ETSI TR 103 761 V1.1.1 (2022-05)

Core Network and Interoperability Testing (INT); Methodologies for E2E Testing & Validation of Vertical Applications over 5G & Beyond networks

Reference

DTR/INT-00170

Keywords

5G, core network, methodology

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Core Network and Interoperability Testing (INT).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

The purpose of the present document is to provide recommendations on methodologies for end-to-end testing and validation of vertical applications over 5G and beyond networks. The present document includes recommendations covering the most aspects of a B5G-app validation framework by providing recommendation on B5G capabilities and enablers, on the testing and validation environment, on involved processes, on the relevant KPI mechanisms and, finally, on the design of vertical applications under test. Such recommendations can be equally applicable to a wide range of industry verticals, application cases and beyond 5G scenarios.

The main value of such end-to-end testing and validation activity is the fact that the vertical application provider can experiment with the 5G and beyond network in order to make business decisions previously to going into commercialization. In this context it is assumed that the subject under test is the application and that the 5G and beyond network setup (as well as the configurations considered in such experimentation) have already been tested and qualified both from functionality and performance perspective.

To that end, the present document provides a survey and review of the existing methodologies for testing and validating vertical applications, leveraging on the experiences gain through several innovation projects. This exercise permits the identification of existing gaps in such methodologies, proposing solutions to cover them. The present document describes Processes, Mechanisms and Strategies involved in the testing and validation of innovative vertical applications enabled by 5G and Beyond networks. The work does not consider any assumption on the specific business or nature of the vertical domain, with the intention to identify common methodologies applicable to as wide a range of vertical domains as possible.

Introduction

Most vertical industries are transforming their processes and innovating their business model, and for that purpose they are actively exploring and adopting a wide range of new technologies. In particular, the adoption of 5G for overcoming limitations and challenges of connectivity and flexibility of other technologies is regarded instrumental for their success. The new 5G landscape of architectures, evolving features and superior performance levels enables possibilities for vertical industries in its digital transformation, and therefore 5G has become a subject of priority focus all along their innovation life-cycle for new applications and solutions, from business opportunity identification to new application's design, solution integration and technical and business validation.

Actually, from early stages of 5G standardization to its regulation, and first commercial deployments, a number of vertical industries have engaged not only on proactive surveillance of 5G technology but also in tight collaborations with Communication Service Providers (CSPs), Telecommunications Equipment Vendors, Academic and Research Institutions and start-ups. That has been a major factor in the steering and shaping of new innovation ecosystems around 5G all around the world, being a remarkable example the one boosted in Europe by the 5G Infrastructure Public Private Partnership (5G PPP) which is a joint initiative between the European Commission and European ICT industry. Virtually all initiatives and projects have been promoted by the 5G PPP. On the one hand, 5G PPP has analysed transformation opportunities of major players of vertical industry sectors, with special attention to end-to-end application requirements and, in turn, their expectations on connectivity and flexibility of the underlying 5G network. On the other hand, 5G PPP has studied and validated the feasibility of 5G architectures and solutions for fulfilling those expectations.

This type of prior-to-commercialization critical validation activities, and that of their implicit challenges motivated, in 2018, the substantiation of large European infrastructures (namely 5G EVE, 5G-VINNI and 5GENESIS projects) for effectively and efficiently host and run the increasing number of 5G-ready application validation activities in the 5G PPP ecosystem. Furthermore, in 2019, the experience and learnings in validation activities translated into the creation of a special 5G PPP workgroup devoted to the applied science of Testing, Measurement and Validation (TMV). The workgroup has been collecting and analysing rich and varied information, from a broad set of projects, on their challenges, approaches, methodologies and tools producing guidelines and recommendations for piloting, adoption, design and execution of validation activities.

Such experience can serve as a good basis for sharing and applying beyond the 5G PPP ecosystem. And given the pace of evolution of 5G towards B5G networks, combined with the intense innovation in vertical applications, a further and careful look into the upcoming challenges for validating applications over B5G networks, seems to be well justified too. So, the ambition of the present document is to leverage the first-hand experience and learnings in 5G PPP, and to assess and provide a basic set of recommendations for crafting effective capabilities, processes and mechanisms for validating vertical applications over 5G and beyond networks.

Despite the reference work being based on initiatives triggered at European level, the outcomes of the present document intend to be globally applicable, without limitation to specific geographic conditions or circumstances. The focus of analysis also sees to that the expectations from the relevant stakeholders involved in the innovation and validation cycle of innovative vertical applications relying on 5G and beyond networks are specified, analysed and supported by the proposed recommendations.

1 Scope

The purpose of the present document is to provide recommendations on methodologies for end-to-end testing and validation of vertical applications over 5G and beyond networks. Such recommendations can be equally applicable to a wide range of industry verticals, application cases and beyond 5G scenarios.

By applying such standard end to end testing and validation methodologies, the vertical application provider would be able to experiment with the 5G and beyond network in a systematic and consistent manner and make informed business decisions upon about further development and commercialization of the features of its application under that rely on 5G technologies and beyond. In this context it is assumed that the subject under test is the application and that the 5G and beyond network set up and configurations considered in such experimentation have already been tested and qualified both from functionality and performance perspective.

Therefore, general (application-independent) testing and characterization of 5G network setups is not in the scope of analysis of the present document. Actually, special emphasis on making clear distinction between network testing and application validation concerns and distinct challenges is secured along the analysis since the processes, mechanisms, tools and strategies typically used for network testing can prove inadequate or misleading for vertical application validation purposes.

The present document provides a survey and review of the existing methodologies for testing and validating vertical applications, to identify existing gaps in such methodologies and propose solutions to cover them. The proposed methodology describes Capabilities, Processes and Mechanisms involved in the testing and validation of innovative vertical applications enabled by 5G and Beyond networks.

No assumptions are made on the specific business or nature of the vertical domain, with the intention to identify common methodologies applicable to as wide a range of vertical domains as possible. The analysis of specific methodologies for applicability limited to specific vertical domains is beyond the scope of the present document.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] 5G PPP Test, Measurement, and KPIs Validation (TMV) Working Group white paper: "Validating 5G Technology Performance Assessing 5G architecture and Application Scenarios", June 2019.

NOTE: Available at https://5g-ppp.eu/wp-content/uploads/2019/06/TMV-White-Paper-V1.1-25062019.pdf.

[i.2] ETSI TS 138 521-3 (V15.4.1): "5G; NR; User Equipment (UE) conformance specification; Radio transmission and reception; Part 3: Range 1 and Range 2 Interworking operation with other radios (3GPP TS 38.521-3 version 15.4.1 Release 15)".

[i.3] ETSI TS 138 521-2 (V15.4.0): "5G; NR; User Equipment (UE) conformance specification; Radio transmission and reception; Part 2: Range 2 standalone (3GPP TS 38.521-2 version 15.4.0 Release 15)".

[i.4] ETSI TS 138 521-1 (V15.3.0): "5G; NR; User Equipment (UE) conformance specification; Radio transmission and reception; Part 1: Range 1 standalone (3GPP TS 38.521-1 version 15.3.0 Release 15)".

[i.5] ETSI TS 138 521-4 (V15.2.0): "5G; NR; User Equipment (UE) conformance specification; Radio transmission and reception; Part 4: Performance (3GPP TS 38.521-4 version 15.2.0 Release 15)".

[i.6] ETSI TS 138 533 (V15.2.0): "5G; NR; User Equipment (UE) conformance specification; Radio Resource Management (RRM) (3GPP TS 38.533 version 15.2.0 Release 15)".

[i.7] ETSI TR 137 901 (V15.1.0): "Universal Mobile Telecommunications System (UMTS); LTE; User Equipment (UE) application layer data throughput performance (3GPP TR 37.901 version 15.1.0 Release 15)".

[i.8] ETSI TR 137 901-5 (V16.3.0): "5G; Study on 5G NR User Equipment (UE) application layer data throughput performance (3GPP TR 37.901-5 version 16.3.0 Release 16)".

[i.9] ETSI TS 128 554 (V15.5.0): "5G; Management and orchestration; 5G end to end Key Performance Indicators (KPI) (3GPP TS 28.554 version 15.5.0 Release 15)".

[i.10] ETSI TS 128 552 (V15.6.0): "5G; Management and orchestration; 5G performance measurements (3GPP TS 28.552 version 15.6.0 Release 15)".

[i.11] 3GPP TS 28.553 (V0.4.0) (July 2018): "Management and orchestration of networks and network slicing; 5G Core Network (5GC) performance measurements and assurance data".

[i.12] ETSI TS 132 425 (V15.3.0): "LTE; Telecommunication management; Performance Management (PM); Performance measurements Evolved Universal Terrestrial Radio Access Network (E-UTRAN) (3GPP TS 32.425 version 15.3.0 Release 15)".

[i.13] ETSI TS 132 450 (V15.1.0): "Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Key Performance Indicators (KPI) for Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Definitions (3GPP TS 32.450 version 15.1.0 Release 15)".

[i.14] Recommendation ITU-T Q.API4TB: "Open APIs for interoperable testbed federations", work in progress, 2021.

[i.15] Recommendation ITU-T Q.3960: "Framework of Internet related performance measurements".

[i.16] Supplement 71 to ITU-T Q-series of Recommendations: "Testing methodologies of Internet related performance measurements including e2e bit rate within the fixed and mobile operator's networks".

[i.17] ETSI TS 103 222-4 (V1.1.1): "Speech and multimedia Transmission Quality (STQ); Reference benchmarking, background traffic profiles and KPIs; Part 4: Reference benchmarking for IPTV, Web TV and RCS-e Video Share".

NOTE: Available at https://www.etsi.org/deliver/etsi_ts/103200_103299/10322204/01.01.01_60/ts_10322204v010101p.pdf.

[i.18] ETSI TS 103 222-3 (V1.1.1): "Speech and multimedia Transmission Quality (STQ); Reference benchmarking, background traffic profiles and KPIs; Part 3: Reference benchmarking, background traffic profiles and KPIs for UMTS and VoLTE".

NOTE: Available at https://www.etsi.org/deliver/etsi_ts/103200_103299/10322203/01.01.01_60/ts_10322203v010101p.pdf.

[i.19] ETSI TS 103 222-2 (V1.1.1): "Speech and multimedia Transmission Quality (STQ); Reference benchmarking, background traffic profiles and KPIs; Part 2: Reference benchmarking and KPIs for High speed internet".

NOTE: Available at https://www.etsi.org/deliver/etsi_ts/103200_103299/10322202/01.01.01_60/ts_10322202v010101p.pdf.

[i.20] ETSI TS 103 195-2 (V1.1.1): "Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture; Part 2: An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management".

NOTE: Available at https://www.etsi.org/deliver/etsi_ts/103100_103199/10319502/01.01.01_60/ts_10319502v010101p.pdf.

[i.21] White Paper No.5 of the ETSI 5G PoC: "Artificial Intelligence (AI) in Test Systems, Testing AI Models and ETSI GANA Model's Cognitive Decision Elements (DEs) via a Generic Test Framework for Testing GANA Multi-Layer Autonomics & their AI Algorithms for Closed-Loop Network Automation".

NOTE: Available at https://intwiki.etsi.org/images/archive/20200527152913%21ETSI_5G_PoC_White_Paper_No_5.pdf.

[i.22] ETSI TC INT/AFI WG PoC (Proof-Of-Concept) Program on 5G Network Slices Creation, Autonomic & Cognitive Management & End-to-End (E2E) Orchestration; with Closed-Loop (Autonomic) Service Assurance of 5G Slices.

NOTE: Available at https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals.

[i.23] ETSI TR 103 748: "INT Artificial Intelligence (AI) in Test Systems and Testing AI models; Use and benefits of AI technologies in Testing,".

NOTE: Available at https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=59455.

[i.24] ETSI TR 103 749: "INT Artificial Intelligence (AI) in Test Systems and Testing AI models; Testing of AI with definition of quality metrics".

NOTE: Available at https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=59456.

[i.25] ETSI TR 103 763: "Core Network and Interoperability Testing (INT); Description of Test Requirements and Approach for E2E Federated Testbeds".

NOTE: Available at https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?WKI_ID=59577.

[i.26] NGMN Alliance: "5G End-to-End Architecture Framework", v3.0.8.

NOTE: Available at https://www.ngmn.org/publications/5g-end-to-end-architecture-framework-v3-0-8.html.

[i.27] 5G-ACIA: "5G for Connected Industries and Automation", Second Edition, White Paper, 5G ACIA, February 2019.

[i.28] 5G-ACIA: "5G for Automation in Industry - Primary use cases, functions and service requirements", White Paper, 5G ACIA, March 2019.

[i.29] 5G-ACIA: "5G Non-Public Networks for Industrial Scenarios", White Paper, 5G ACIA, July 2019.

[i.30] 5G-ACIA: "Key 5G Use Cases and Requirements - From the Viewpoint of Operational Technology Providers", White Paper, 5G ACIA, May 2020.

[i.31] 5G-ACIA: "Integration of 5G with Time-Sensitive Networking for Industrial Communications", White Paper, 5G-ACIA, January 2021.

[i.32] 5G PPP Phase 1 projects.

NOTE: Available at https://5g-ppp.eu/5g-ppp-phase-1-projects/.

[i.33] 5G PPP Phase 2 projects.

NOTE: Available at https://5g-ppp.eu/5g-ppp-phase-2-projects/.

[i.34] 5G PPP Phase 3 projects.

NOTE: Available at https://5g-ppp.eu/5g-ppp-phase-3-projects/.

[i.35] 5G-EVE Project.

NOTE: Available at https://www.5g-eve.eu/.

[i.36] 5GENESIS Project.

NOTE: Available at https://5genesis.eu/.

[i.37] 5G-VINNI Project.

NOTE: Available at https://www.5g-vinni.eu/.

[i.38] OSM.

NOTE: Available at https://osm.etsi.org/.

[i.39] OPNFV.

NOTE: Available at https://www.opnfv.org/.

[i.40] ONAP.

NOTE: Available at https://www.onap.org/.

[i.41] ETSI ZSM.

NOTE: Available at https://www.etsi.org/committee/zsm.

[i.42] 5GPP Test, Measurement, and KPIs Validation (TMV) Working Group white paper: "Service performance measurement methods over 5G experimental networks", May 2021.

NOTE: Available at https://5g-ppp.eu/wp-content/uploads/2021/06/Service-performance-measurement-methods-over-5G-experimental-networks_08052021-Final.pdf.

[i.43] ETSI TS 138 101-1 (V15.13.0): "5G; NR; User Equipment (UE) radio transmission and reception; Part 1: Range 1 Standalone (3GPP TS 38.101-1 version 15.13.0 Release 15)".

[i.44] ETSI TS 138 101-2 (V15.13.0): "5G; NR; User Equipment (UE) radio transmission and reception; Part 2: Range 2 Standalone (3GPP TS 38.101-2 version 15.13.0 Release 15)".

[i.45] ETSI TS 138 211 (V15.9.0): "5G; NR; Physical channels and modulation (3GPP TS 38.211 version 15.9.0 Release 15)".

[i.46] ETSI TS 128 530 (V15.1.0): "5G; Management and orchestration; Concepts, use cases and requirements (3GPP TS 28.530 version 15.1.0 Release 15)".

[i.47] Recommendation ITU-T Y.3100 (09/2017): "Terms and definitions for IMT-2020 network".

[i.48] ETSI TS 123 501 (V15.4.0): "5G; System Architecture for the 5G System (3GPP TS 23.501 version 15.4.0 Release 15)".

[i.49] ETSI MEC.

NOTE: Available at https://www.etsi.org/committee/mec.

[i.50] Linux Foundation Akraino.

NOTE: Available at https://www.lfedge.org/projects/akraino/.

[i.51] Openess.

NOTE: Available at https://www.openness.org.

[i.52] ETSI NFV.

NOTE: Available at https://www.etsi.org/committee/nfv.

[i.53] iPerf.

NOTE: Available at https://iperf.fr/.

[i.54] ETSI GS NFV-SOL 005 (V2.7.1): "Network Functions Virtualisation (NFV) Release 2; Protocols and Data Models; RESTful protocols specification for the Os-Ma-nfvo Reference Point".

[i.55] ETSI TS 129 522 (V15.8.0): "5G; 5G System; Network Exposure Function Northbound APIs; Stage 3 (3GPP TS 29.522 version 15.8.0 Release 15)".

[i.56] ETSI TS 129 222 (V15.9.0): "5G; LTE; Common API Framework for 3GPP Northbound APIs (3GPP TS 29.222 version 15.9.0 Release 15)".

[i.57] 3GPP TR 23.758 (V17.0.0): "Study on application architecture for enabling Edge Applications".

[i.58] ETSI GS MEC 012 (V2.1.1): "Multi-access Edge Computing (MEC); Radio Network Information API".

[i.59] ETSI GS MEC-DEC 032-3 (V2.1.1): "Multi-access Edge Computing (MEC); API Conformance Test Specification; Part 3: Abstract Test Suite (ATS)".

[i.60] Report ITU-R M.2410-0 (11/2017): "Minimum requirements related to technical performance for IMT-2020 radio interface(s)".

[i.61] 5G-VINNI Deliverable D4.1: "Initial report on test-plan creation and methodology, and development of test orchestration framework", July 2019.

NOTE: Available at https://zenodo.org/record/3345626.

[i.62] ETSI TR 103 747: "Core Network and Interoperability Testing (INT/ WG AFI); Federated GANA Knowledge Planes (KPs) for Multi-Domain Autonomic Management & Control (AMC) of Slices in the NGMN(R) 5G End-to-End Architecture Framework".

[i.63] Recommendation ITU-T Y.3324: "Requirements and architectural framework for autonomic management and control of IMT-2020 networks".

[i.64] ETSI TR 103 473: "Evolution of management towards Autonomic Future Internet (AFI); Autonomicity and Self-Management in the Broadband Forum (BBF) Architectures".

[i.65] ETSI TR 103 404: "Network Technologies (NTECH); Autonomic network engineering for the self-managing Future Internet (AFI); Autonomicity and Self-Management in the Backhaul and Core network parts of the 3GPP Architecture".

[i.66] ETSI TS 132 404: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Performance Management (PM); Performance measurements; Definitions and template (3GPP TS 32.404)".

[i.67] 3GPP TR 22.804: "Study on Communication for Automation in Vertical domains (CAV)".

[i.68] ETSI TS 122 261: "5G; Service requirements for the 5G system (3GPP TS 22.261)".

[i.69] ETSI TS 122 104: "5G; Service requirements for cyber-physical control applications in vertical domains (3GPP TS 22.104)".

[i.70] 3GPP TR 22.830: "Study on business role models for network slicing".

[i.71] 3GPP TS 28.531 (V16.0.0): "Management and orchestration; Provisioning".

[i.72] White paper 5G PPP: "5G network support of vertical industries in the 5G Public-Private Partnership ecosystem".

NOTE: Available at https://5g-ppp.eu/wp-content/uploads/2020/02/Vertical-industries-in-the-5G-PPP.pdf.

[i.73] ETSI GANA Model in 5G Network Slicing PoC White Paper #4: "ETSI GANA as Multi-Layer Artificial Intelligence (AI) Framework for Implementing AI Models for Autonomic Management & Control (AMC) of Networks and Services; and Intent-Based Networking (IBN) via GANA Knowledge Planes (KPs)".

NOTE: Available at https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals.

[i.74] TMForum IG1127: "End-to-end Virtualization Management: Impact on E2E Service Assurance and SLA Management for Hybrid Networks".

[i.75] ETSI GANA Model in 5G Network Slicing PoC White Paper #3: "Programmable Traffic Monitoring Fabrics that enable On-Demand Monitoring and Feeding of Knowledge into the ETSI GANA Knowledge Plane for Autonomic Service Assurance of 5G Network Slices; and Orchestrated Service Monitoring in NFV/Clouds".

NOTE: Available at https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals.

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the following terms apply:

5G PPP Facility: each one of the testing sites which form an ICT-17 5G validation platform. Thus, each platform has several facilities in different geographical locations

5G Platform: ICT-17 platforms funded by EU Commission: 5GEVE, 5GVINNI and 5Genesis

5GS observation points: observation points located on interfaces within the 5G System, including the 5G Radio, Edge, Transport and 5G Packet Core

APP E2E observation points: observation points located on the hardware and software application or services that the vertical controls or owns

B5G Network: beyond 5G Networks are networks built with technology that is specified by future releases of 3GPP after release 17, and are planned to be introduced starting 2025

Communication Service Provider (CSP): company that offers communication services, typically, Network Operators offering Public Mobile services communications

experiment: running of a specific process that includes End user devices, 5G Network components and Vertical application to discover KPI values that are not know in advance

experiment blueprint: set of composed actions including end user devices, network components, vertical application, test cases, measurements, and KPIs that can be introduced in a 5G Facility to characterize the behaviour of the system under specific configuration

Management and Orchestration (MANO): framework developed by a working group of the same name within the European Telecommunications Standards Institute (ETSI) Industry Specification Group for NFV (ETSI ISG NFV)

NOTE: It is the ETSI-defined framework for the management and orchestration of all resources in a virtualized data centre including computer, networking, storage, and Virtual Machine (VM) and Container resources.

(Network) Monitoring: monitoring is a computer network's systematic effort to detect slow or failing network components, such as overloaded or crashed/frozen servers, failing routers, failed switches or other problematic devices

NOTE: In the event of a network failure or similar outage, the network monitoring system alerts the network administrator. Network monitoring is a subset of network management.

performance: in the context of Networking, analysis and review of collective network statistics, to define the quality of services offered by the system considering end to end interactions between end user devices and vertical applications

subject under test: artefact that is being evaluated for testing purposes

NOTE: Typically, in the context of 5G Platforms, the subject under test matches with Vertical Applications.

test: process of validating either a functional or non-functional behavior of a system (e.g. device, software component, etc.)

NOTE: In the context of the present document, a Test consist in running a specific process that includes End user devices, 5G Network components and Vertical Application to obtain KPI values and verify that obtained values fit in predefined thresholds.

T&M methodologies: methods, rules and processes required to test and measure results of these tests

Test, Measurement, and KPIs Validation (TMV) Working Group: one of the working groups in 5G PPP whose objective is to bring together the projects within the 5G PPP that have common interest in the development of Test and Measurement methods, test cases, procedures and KPI validation

Test as a Service (TaaS): automation and interfacing layer that allows to connect all the Test & Measurement tools needed for validating and verifying a system, from the individual components up to the E2E service

NOTE: It speeds up repeating tests and validating proper behaviour of a system after introducing changes.

testing: process of evaluating a system or its component(s) with the intent to find whether it satisfies the specified requirements

validation cycle: collaborative process between Vertical industries and Communication Service Providers sharing objectives, timelines, outcomes and learnings to guarantee proper integration between Vertical applications and Communications networks

vertical application: software program that performs specific data processing related to specific domain. Examples of vertical are Factory of the Future (FoF), Unmanned Aerial Systems (UAS), Vehicle to Infrastructure Communications (V2X) and Edge Applications (EDGEAPP)

vertical KPI model: relationship between service KPIs as defined by the vertical and network KPIs enforced by the provider

NOTE: It refers to a model which represents the influence of network service KPIs on the Vertical-level KPIs. It can be presented in a tabular structure (mapping) or in a more complex form.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

5G PPP	5G Infrastructure Public Private Partnership
5G-RAN	5G-Radio Access Network
5GS	5G System
AFI	Autonomic Management and Control Intelligence for Self-Managed Fixed & Mobile Integrated Networks
AI	Artificial Intelligence
AIM	Automated Intelligent Management
AMC	Autonomic Management and Control
AMF	Access and Mobility Management Function
AN	Access Network
API	Application Programming Interface
APP	Application
B5G	Beyond 5G
BBF	Broadband Forum
CC	Component Carriers
CI/CD	Continuous Integration/Continuous Deployment

CP	Cyclic Prefix
CPU	Central Processing Unit
CSC	Communication Service Customer
CSP	Communication Service Provider
CSV	Comma Separated Value
DL	Downlink
DMRS	Demodulation Reference Signal
DTR	Draft Technical Report
E2E	End-to-End
EEM	Experiment Execution Manager
ELCM	Experiment LifeCycle Manager
eMBB	enhanced Mobile Broadband
EU	European Union
FDD	Frequency Division Duplex
F-MBTS	Federated Model-Based-Translation Service
GANA	Generic Autonomic Networking Architecture
GPU	Graphics Processing Unit
GSMA	Groupe Speciale Mobile Association
GST	Generic Slice Template
GUI	Graphical User Interface
HTML	Hypertext Markup Language
IBN	Intent-based Networking
ICT	Information and Communication Technology
ID	Identifier
IG	Introductory Guide
IMT	International Mobile Telecommunications
IoT	Internet of Things
IP	Internet Protocol
IPTV	Internet Protocol television
IT	Information Technology
JSON	JavaScript Object Notation
KP	Key Performance
KPI	Key Performance Indicator
LPWA	Low Power Wide Area
MANO	Management and Orchestration
MBTS	Model-Based-Translation Service
MEC	Multi-access Edge Computing
MIMO	Multiple-input multiple-output
MIoT	Massive Internet of Things
ML	Machine Learning
mMTC	massive Machine Type Communications
MNO	Mobile Network Operator
MU	Multi-User
NEF	Network Exposure Function
NetOp	Network Operation
NFV	Network Function Virtualization
NFVI	NFV Infrasrtucture
NGI	Next Generation Internet
NGMN	Next Generation Mobile Networks
NGRAN	Next Geenration Radio Access Network
NG-RAN	Next Geenration-Radio Access Network
NS	Network Slice
NSA	Non-standalone
NSD	Network Service Descriptor
NSI	Network Slice Instance
NSMF	Network Slice Management Function
NSSF	Network Slice Selection Function
NST	Network Slice Template
OAI	Open Air Interface
OAM	Operation And Maintenance
OFDM	Orthogonal Frequency Division Multiplexing
OLA	Operational Level Agreement

ONAP	Open Network Automation Platform
ONIX	Overlay Network system of information servers for Information eXchange
OPEX	Operational Expenditure
OPNFV	Open Platform for Network Function. Virtualization
OSM	Open Source MANO
PDF	Portable Document Format
PDSCH	Physical Downlink Shared Channel
PHP	Hypertext Preprocessor
PLMN	Public Land Mobile Network
PNG	Portable Network Graphics
PoC	Proof of Concept
QAM	Quadrature Amplitude Modulation
QoE	Quality of Experience
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying
R&D	Research and Development
RAN	Radio Access Network
RAV	Real-Time Analytics and Validation
RCA	Root Cause Analysis
RCS	Rich Communication Services
REST	REpresentational State Transfer
RRC	Radio Resource Control
RRM	Radio Resource Management
RT	Real Time
RTC	Run-Time Configuration
RTT	Round Trip Time
SA	Standalone
SBA	Service-Based Architecture
SDK	Software Development Kit
SDN	Software Defined Networks
SDO	Standardization Development Organization
SLA	Service Level Agreement
SLO	Service Level Objectives
SLS	Service Level Specification
SME	Small and Medium Enterprise
SMF	Session Management Function
SQL	Structured Query Language
STQ	Speech and multimedia Transmission Quality
SU	Single User
SUT	Subject Under Test
T&M	Test and Measurement
TaaS	Test as a Service
TAP	Testing Automation Platform
TDD	Time Division Duplex
TIM	Telecom Italia
TMV	Testing, Measurement and Validation
TV	Television
UE	User Equipment
UI	User Interface
UL	Uplink
UPF	User Plane Function
URL	Uniform Resource Locator
URLLC	Ultra-Reliable Low Latency Communications
UTRAN	UMTS Terrestrial Radio Access Network
V2X	Vehicle to Infrastructure Communications
VF	Virtual Function
VIM	Virtual Infrastructure Manager
VM	Virtual Machine
VNF	Virtual Network Functions
VNFD	VNF Descriptor
VoLTE	Voice over LTE
VPN	Virtual Private Network

ZSM Zero touch network & Service Management

4 Methodologies for Testing and Validation of Vertical Applications over 5G & Beyond Networks

4.1 Motivation and expectations

4.1.1 Stakeholders

The key stakeholders in the validation of innovative vertical applications designed to leverage Beyond 5G (B5G) networks are the vertical firm/business and its partner the Communication Service Provider (CSP), facilitating the access and use of the required platform, capabilities, processes and tools for designing, planning executing the validation tests agreed upon them.

It could be argued, as it has been actually the case in some early exploratory activities and projects, that the vertical business could partner, instead of with a CSP, with, for instance, a research institution acting as a CSP-independent validation platform/service provider. For the sake of simplicity, and without compromising the generalization of conclusions and recommendations in the present document, the CSP term is used for designating the actor partnering with the vertical in their validation endeavour.

Similarly, even though the innovation ecosystems of Verticals and CSPs extend to a number of players supporting each or both of them (such as Independent Software Vendors (ISVs), Telecom Equipment Manufacturers, Chipset and Device Manufacturers, etc.), also for the sake of simplicity, when referring to the needs and expectations for recommendations to Verticals and CSPs, it is implied that they refer and apply to both but also to the respective ecosystems supporting them.

The remarkable aspect, when it comes to the stakeholders involved in validation activities, is that Verticals and CSPs are trusted partners to each other along the validation cycle, with both sharing objectives, timelines, outcomes and learnings. In the validation cycle the vertical firm will develop assurance on the true potential of 5G and Beyond related technologies for expanding their solutions or even innovating their business model, being launched with the CSP partnering during the validation and/or as with any other CSPs worldwide.

The CSPs will gain insight into the type of vertical applications as potential providers of such solutions, which will help them make decisions on whether and how to evolve their portfolio to support verticals with similar demands on 5G network services. So, that binding between these two players (who otherwise could be regarded in a very simplistic way as customer and supplier) proves to be the most fruitful approach for both of them as a way to expand their knowledge, technology and portfolio, in order to seize new business opportunities in the market.

The vertical innovation cycle will require extensive validation and experimental testing in order to ensure the proper and correct functionality and performance of vertical applications when using 5G and Beyond technologies. The experimental facilities should perfectly mimic the conditions and configuration to be found in production networks to verify whether the vertical application performs as expected. It is also required to work in a formalization process including common methodological approaches and information processes for experimentation, as reflected in [i.1].

Considering the actual expectations, shared by the above introduced Stakeholders, on the validation of vertical applications over B5G networks, the most relevant ones are:

i) Application-network interoperability verification.

ii) Application end-to-end performance evaluation.

iii) Network technology suitability assessment for serving the application under test.

iv) Application deployment model optimization.

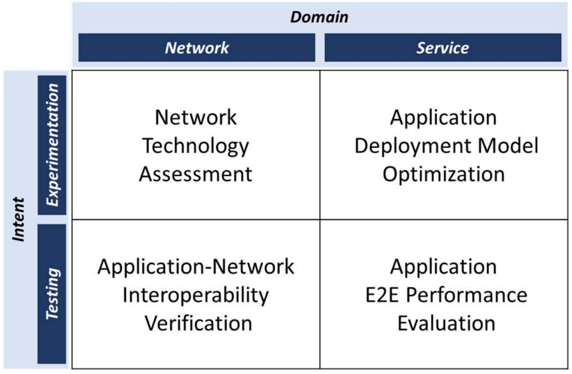

Figure 1: Stakeholders expectations

Therefore, as illustrated in Figure 1, both testing and experimentation complementary intents are considered, and also both network and service dimensions are taken into account. Altogether, this provides a broad field of analysis for the present document. Each of these expectation is analysed in following sections, outlining the distinct pursued goals and benefits, as well as the accepted conditions and limitations applying.

4.1.2 Expectations

4.1.2.1 Application-network interoperability verification

The goal of this expectation is to validate the technical compatibility of the vertical application with a specific 5G setup and configuration, which is representative of the *a priori* type of targeted 5G deployment environment assumed for taking the vertical application to operation.
The main benefit is to secure an early and agile adaptation of the vertical application to 5G, so that its design and architecture is compatible with one or several flavours of 5G. This is a precondition to achieve the remaining validation goals.

These tests are normally designed to assess and validate basic and flawless interoperability (i.e. application deployment, network connectivity, proper interworking with devices, etc.). The formal validation process bases on clear pre-conditions and post-conditions, defined beforehand, and some specific functional and performance KPIs can be defined, obtained and used along the validation process. However, this is only a first step before further validating, extensively, the functionality and the performance of the vertical application.

4.1.2.2 Application end-to-end performance assurance and characterization

Two staged levels of ambition are considered for this expectation: application viability assurance, and application performance characterization.

The first goal (application viability assurance) is to determine *whether or not* a specific 5G setup and configuration delivers the expected 5G services and performance making it possible for the application to deliver its service functionality with the expected end-to-end performance.
The main benefit is the early confirmation of the viability of the vertical application over 5G (even if only assured for the specific network setup and configuration addressed). These test are normally limited to low-scale application scenarios, in order to just check, as early as possible, that the behaviour and performance of the vertical application can indeed meet the set expectations, or else get knowledge that it does not.

The second goal (application performance characterization) is to determine and characterize *to what extent* a specific 5G setup and configuration delivers the expected 5G services and performance, making it possible for the application to behave as expected both in terms of functionality and end-to-end performance in large-scale operations-comparable scenarios.
The main benefit is the full characterization (and validation) of the application performance levels over the selected 5G setup and configuration. These tests consider large-scale application scenarios and a broad range of operational conditions. Each test is limited in practice, however, to a fixed selected type of 5G network setup and configuration

At this stage an extensive use of hand-shaken well-defined beforehand focused Vertical and Network KPIs is expected for defining the post-conditions to be objectively and quantitatively assessed.

4.1.2.3 Network technology assessment and selection for serving the application under test

The goal in this case is to determine and characterize to what extent a variety of 5G setups and configurations deliver (or not) the expected 5G services and performance making it possible for the application to behave as expected both in terms of functionality and end-to-end performance, from low-scale testing to large-scale operations-comparable application scenarios validation. The technology capabilities and setups under interest may be evaluated and selected as viable beforehand.

The main benefit is the identification and comparative analysis of 5G setups as well as configurations that may cope with the needs of the vertical applications under test, so that the vertical application can deliver the functionality and end-to-end performance expected.

For this expectation to be met, the post-conditions to be assessed are expressed through a subset from the set of focused KPIs at both vertical and network levels, relevant to the *comparison* of technologies. And beyond that specific technical KPIs analysis, each stakeholder defines a complementary set of business/economic KPIs related to the technologies involved, in order to achieve a comprehensive techno-economic assessment and rating of alternatives.

4.1.2.4 Application deployment model evaluation and optimization

The goal is to get an insight on the influence of alternative connectivity deployment models of the vertical application over a selected 5G setup and configuration. The application connectivity deployment models of interest may be evaluated and selected as viable beforehand.
The main benefit is the identification and comparative analysis of application connectivity deployment models with regards to their delivery of expected functionality and end-to-end performance.

For this expectation to be met, the post-conditions to be assessed are expressed through a subset from the set of focused KPIs at both vertical and network levels, relevant to the *comparison* of connectivity deployment models. And further business/economic KPIs are considered by each stakeholder in order to achieve a comprehensive assessment and rating of alternatives.

4.1.3 Key Variables: Inputs and Outputs

Now, after having described the basic principles and objectives for the set of complementary validation dimensions, it is important to also reflect on, and formally model, the space of variables (inputs and outputs) involved in a generic validation process. An enumeration of those variables for a complete and robust validation system should include, at least the following:

a) The vertical application itself (i.e. the Subject Under Test, SUT).

b) The set of application-specific KPIs, with the definition and observation points, describing its expected end-to-end performance (i.e. the indicators the application is validated for).

c) The range of application-specific operational and environmental conditions (i.e. the conditions the application is validated against).

d) The 5G setup and configuration (i.e. the mobile and complementary network infrastructure where the application is validated).

e) The application deployment model over the 5G network (i.e. the options of deployment against which the application is validated).

4.1.4 The Validation Cycle

The whole validation cycle has to pivot around the vertical application, being the Subject Under Test, thus the key and common input to all sorts of validation activities.

First of all, validation activities, essentially, are expected to deliver formal evaluation results of such tested application against the testing scenarios defined by the choices made for all the other variables enumerated above, and with a high level of confidence.

Then, the validation cycle is expected to determine not only that formal evaluation for closed scenarios, but actually help both the vertical and the communication service provider to identify the scenarios for which the application can deliver the functionality and end-to-end performance that qualifies for its business purpose.

Next, the insights obtained along the validation cycle can trigger a new iteration of the innovation cycle at both vertical and communication service provider sides for either evolving, enhancing or simply tuning, respectively, their application and network, in order to better meet the demands of their respective businesses.

And, finally, an eventual further validation cycle can be started for getting further assurance and gaining new insights. As a result, the way is paved for both verticals and CSPs to make technology and business decisions based on consistent facts and learnings.

So, in summary, that is the type of high-level validation cycle pursued for enabling verticals and communication service providers to share objectives, collaborate in validation activities, converge to conclusions, and become ready for making well-informed technology and business decisions. The requirements that such ambitious validation cycle imposes on the capabilities to be deployed, on the processes to be enabled, on the basic mechanisms to operationalize them, and on the strategies for increasing the level of confidence in the validation cycle.

4.1.5 The Validation Framework

The main system proposed for meeting the expectations of the relevant Stakeholders on the validation cycle of vertical applications over B5G networks is the *B5G-app Validation Framework (B5G-VF)*. The B5G-app Validation Framework a composite system of capabilities, processes, services and tools operated by CSPs and meant to be used by partner/engaged Verticals. The recommendations for building out, operating and using B5G-app Validation Frameworks is the ultimate goal of the present document.

4.2 State of the Art survey

4.2.1 Standards of relevance

4.2.1.1 3GPP

Over the years, 3GPP has defined test cases for multiple technologies, including UMTS (3G), LTE (4G) and now also 5G NR.

There are multiple test specifications addressing different aspects and uses cases of the technology. 3GPP details in different documents how a 5G NR conformance test device should be verified from an RF point of view, characterizing both UE transmitter and received parameters such as maximum transmit power, receiver sensitivity or spurious emissions.

Non-Standalone ETSI TS 138 521-3 [i.2] scenarios combining NR and LTE cells and also Standalone configurations are covered, addressing both mmWave ETSI TS 138 521-2 [i.3] and sub-6 GHz scenarios ETSI TS 138 521-1 [i.4]. Performance aspects such as demodulation under different propagation and SNR conditions, are defined in [i.5]. ETSI TS 138 521-1 [i.4], ETSI TS 138 521-2 [i.3] and ETSI TS 138 521-3 [i.2] specify the testing of the involved 5G NR protocols. Moreover ETSI TS 138 533 [i.6] specifies RRM (Radio Resource Management) test cases, including reporting of power and quality measurements, Handover latency as examples.

ETSI TR 137 901 [i.7] was the only specification from 3GPP that covers testing at the application layer. In particular, it 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. On parallel, 3GPP has also detected a strong demand UE Application- Layer Data Throughput Measurements, triggering a Study on 5G NR User Equipment (UE) application layer data throughput performance as an evolution of ETSI TR 137 901 [i.7], progressing as ETSI TR 137 901-5 [i.8].

ETSI TS 128 554 [i.9] is focused on 5G KPIs and network slicing. It specifies a KPI definition template to allow categorization of the KPIs and the methods, tools, and calculations that are used in order to measure and validate these KPIs.

3GPP specifications provide two more documents that focus on 5G performance measurements ETSI TS 128 552 [i.10] and 5G Core Network (5GC) performance measurements and assurance data 3GPP TS 28.553 [i.11] (this specification was finally withdrawn). The first document provides specifications for the performance measurements of 5G networks including network slicing. Performance measurements for NGRAN as well as for 5GC are defined in the present 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 ETSI TS 132 425 [i.12] and related KPIs in 3GPP TS 32.450 [i.13]). The performance measurements are defined based on the measurement template as described in ETSI TS 132 404 [i.66]. 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 each 5G component separately (i.e. 5GC or 5G-RAN) while ETSI TS 128 554 [i.9] focuses more on end-to-end KPI validation.

All these 3GPP specifications are focused on certain components of the network, even on end-to-end KPIs, however there is a gap in all of them. Those which cover end-to-end KPIs, such as ETSI TS 128 554 [i.9], do not provide the test sequence for executing the test from the point of view of verticals, which, at the end are the "users" of network and need to verify that their services perform accordingly to SLA agreed with their customers. Moreover, the definition of end-to-end scenarios for reproducing the scenarios where the verticals are going to deploy their services is missing. Network conditions (scenarios) can impact heavily into the measurements, these conditions have to be well defined to be able to contextualize and compare the obtained results.

Furthermore, 3GPP works also on standardization activity for Verticals, especially in the industrial domain. In the mandate of 3GPP SA1, requirements that are relevant to Verticals (Factory and Process Automation use cases and related Performance) have been introduced since Release 16 by publishing Technical Reports (TRs) first as usual. The targeted scope of the Technical Reports was "Study on Communication for Automation in Vertical Domains" (3GPP TR 22.804 [i.67]). The second step was the publication of normative specifications in this requirements space (Technical Specifications series: TSs) such as "Service requirements for the 5G system" (ETSI TS 122 261 [i.68]), "Service requirements for cyber-physical control applications in vertical domains" (ETSI TS 122 104 [i.69]) , "Feasibility Study on Business Role Models for Network Slicing" (3GPP TR 22.830 [i.70]) and others. Improvements and enhancements are addressed in Releases 17 and 18.

4.2.1.2 ITU

The current work in ITU-T Q.API4TB [i.14] is addressing the subject of testbeds Federations for 5G and Beyond. As reflected on it, over the years it has increasingly been experienced that isolated standalone testbeds are not sufficient to test and trial out certain technology uses cases because the use cases rather require the use of components and resources located in various testbeds with each testbed bringing in the missing/required features and assets to complete the use-case. On the other hand, new ICT technologies, networks and vertical applications are becoming increasingly complex to test by simply using standalone testbeds. The expectation is for federated testbeds to bring sustainability in fostering environments for quick innovations and testing of complex technologies and use cases, and for enabling quicker time to market for products and services. To be able to test various vertical applications that require testbeds federations in order to test them and measure various KPIs, the ITU-T is currently developing ITU-T Q.API4TB which is expected to guide the development of testbeds that can be federated and can be used for such testing.

In addition to that, some other ITU relevant documents can be found in Recommendation ITU-T Q.3960 [i.15] which describes a framework for Internet related performance measurements which can be established at the national or international level, providing customers of the existing public telecom networks the possibility to estimate the access related performance.
In Supplement 71 [i.16] can be found the testing procedures of data transmission speed within the fixed and mobile operators' networks. The methodology is based on the concept of the Recommendation ITU-T Q.3960 [i.15].

4.2.1.3 ETSI

A number of initiatives carried out in ETSI can be mentioned.

Over the years, ETSI Speech and Transmission Quality (STQ) has defined KPIs for Transmission requirements for Fixed and Mobile from QoS as perceived by the User. There are multiple specifications addressing different aspects and uses cases of the technology. Relevant document can be found in ETSI TS 103 222-4 [i.17], ETSI TS 103 222-3 [i.18], ETSI TS 103 222-2 [i.19] where a multi-part deliverable cover the Reference benchmarking, background traffic profiles and KPIs for ETSI TS 103 222-2 [i.19], ETSI TS 103 222-3 [i.18] and ETSI TS 103 222-4 [i.17].

Besides that, the various types of requirements imposed on 5G and Beyond Networks by Vertical Applications pertaining to optimized E2E latency, resilience and survivability of network services, closed-loop service and security assurance to guarantee E2E QoS and Security SLAs without need for human interventions, call for the 5G and Beyond Networks to operate on the basis of the Autonomic Management and Control (AMC) Paradigm as described in ETSI White Paper No.16, ETSI TS 103 195-2 [i.20], ETSI TR 103 747 [i.62] and White Papers published by the ETSI 5G PoC Project [i.22]. The Subgroup (Working Group) of the ETSI TC INT, namely Autonomic Management and Control (AMC) Intelligence for Self-Managed Fixed & Mobile Integrated Networks (AFI) WG is producing standards for AMC in various network architectures and their associated management and control architectures (including 5G E2E architectures) for Autonomic/Autonomous Networks (ANs). The ETSI TS 103 195-2 [i.20] is the de-facto standard for AMC, as defined by the Generic Autonomic Networking Architecture (GANA) reference Model for Autonomic Networking, Cognitive Networking and Self-Management of Networks and Services, the model specified in ETSI TS 103 195-2 [i.20]. There is other work in ETSI TC INT AFI WG on GANA instantiations onto various network architectures and their associated management and control architectures.

The ETSI GANA multi-layer autonomics and multi-layer AI model for AMC is enabler Autonomic/Autonomous Networks (ANs). As described in ITU-T Q.API4TB [i.14]: Autonomic/Autonomous Networks (ANs), powered by the GANA AMC (Autonomic Management and Control) paradigm (ETSI TS 103 195-2 [i.20], Recommendation ITU-T Y.3324 [i.63], NGMN 5G End-to-End Architecture Framework v3.0.8 [i.26] and employing multi-layer autonomics and multi-layer AI, are expected to be driven by so called Knowledge Planes (KPs) Platforms as discussed in ETSI TS 103 195-2 [i.20], ETSI TR 103 747 [i.62], ETSI TR 103 473 [i.64] (V1.1.2), ETSI TR 103 404 [i.65], NGMN 5G End-to-End Architecture Framework v3.0.8, BroadBand Forum (BBF's) AIM (Automated Intelligent Management) Framework, Recommendation ITU-T Y.3324 [i.63]. Standards for Knowledge Plane (KP) Platforms for ANs now exist, with the main standard being the ETSI GANA (Generic Autonomic Network Architecture) Knowledge Plane (KP) concept specified in ETSI TS 103 195-2 [i.20]. When ANs, as represented by either individual network segments of a CSP (e.g. RAN, Edge Cloud, Transport, Core network) or the whole end-to-end CSP's networks, need to be collaboratively interworked in any beneficial form, this should be achieved by having the KP Platforms that are responsible for AMC of specific network segment as individual ANs communicate with each other in the form of KPs Federations (more details can be found on this subject in ETSI TS 103 195-2 [i.20] and in White Papers available at https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals).

Therefore, there are two kinds of Standards being developed in ETSI TC INT that are relevant to Vertical Applications over Autonomic/Autonomous 5G and Beyond Networks and Testing:

1) Standards for Testing AI Models (including AI Models of GANA Cognitive Decision-making-Elements (DEs) for AMC in 5G and Beyond Networks (ETSI TR 103 748 [i.23], ETSI TR 103 749 [i.24], ETSI TR 103 763 [i.25]). ETSI 5G PoC White Paper No.5 [i.21] describes a way to test GANA cognitive indirectly by measuring network service performance KPIs (e.g. KPIs of relevance to network services like those required by Vertical Applications) with the GANA DEs configured to operate in "open-loop mode" such that the DEs do not directly perform actions on the network, and then repeating the tests and measuring the KPIs with the GANA DEs configured to operate in "closed-loop mode". The analysis and comparison (the difference) of the KPIs data and other changes effected by the DEs in both contexts (open-loop versus closed-loop modes) provide an assessment of the impact or value of DE autonomics for the network.

2) Testing of ANs (e.g. GANA based 5G Autonomic/Autonomous Networks) as Use Case for Standards for Testbeds Federations for 5G and Beyond being developed in ITU-T SG11. Through Federation, knowledge exchange and transfer, meta-data, events, triggers, synchronization and coordination messages, and many other forms of information are communicated in a collaborative fashion by the KP Platforms to achieve E2E AMC operations such as E2E Self-Optimization of AN resources, Self-/Protection and Self-/Defence against detected security attacks, threats and risks to address security challenges that have impact on various network domains. Testing ANs as represented by individual network segments and their associated their KP level autonomics ("slow control-loops") and autonomics ("fast control-loops") introduced in the underlying infrastructure that constitute a specific network segment requires Federated Testbeds that emulate the ANs' compositions and targeted interworking of their two levels of autonomics. There is ongoing work in an ETSI Work Item (WI) for documenting and specifying a Use Case of Federated Testbeds and the Instantiation of the Reference Model for Federated Testbeds being standardized by ITU (ITU-T Q.API4TB [i.14]) with respect to Testing GANA KP Platforms for E2E AMC across multiple domains (network segments and administrative inter CSP domains). This ETSI WI can be found at ETSI TR 103 763 [i.25]. ITU-T Q.API4TB [i.14] provides more details on how Federated Testbeds built on the basis of the Reference Model for Testbeds Federations being standardized jointly by ITU-T SG11 and ETSI TC INT can play a very important role in Testing the emerging Autonomic/Autonomous Networks (ANs) technologies.

4.2.2 Industry Alliance in the Vertical ecosystem: 5G ACIA

Quoting 5G-ACIA (5G Alliance for Connected Industries and Automation), 5G-ACIA was established to serve as the central and global forum for addressing, discussing, and evaluating relevant technical, regulatory, and business aspects with respect to 5G for the industrial domain. The goal is to ensure the best possible applicability of 5G technology and 5G networks to the industrial domain.

This is the reason why, 5G-ACIA involves also Telco Vendors, MNOs (Mobile Network Operators), Chips Manufactures and other stakeholders and relies and leverages 3GPP SA1 assets (TRs and TSs) but also other SDOs work. The main 5G-ACIA published documents are: "5G for Connected Industries and Automation" [i.27], "5G for Automation in Industry - Primary use cases, functions and service requirements" [i.28], "5G Non-Public Networks for Industrial Scenarios" [i.29], "Key 5G Use Cases and Requirements - From the Viewpoint of Operational Technology Providers" [i.30], and "Integration of 5G with Time-Sensitive Networking for Industrial Communications" [i.31].

4.2.3 The European Commission (EC) 5G ICT projects as state of the art

4.2.3.1 Introduction to 5G PPP program

Whilst 5G is being currently deployed mainly in commercial networks for enhanced Mobile Broadband (eMBB) services, extended pilot trials are being executed around the world to validate 5G also for other vertical use cases. Such trials cover multiple vertical domains, like autonomous driving, smart factories, healthcare, media, energy, etc.

Some remarkable examples of state-of-the-art solutions can be found in the 5G Infrastructure Public Private Partnership (5G PPP) initiative. No other state of the art was identified at the time of writing the present document, even though the present document does not preclude other existing solutions been representative for the purpose of the present document.

5G PPP is a joint initiative between the European Commission (EC) and the European ICT industry (ICT manufacturers, telecommunications operators, service providers, SMEs and researcher Institutions). 5G PPP aims to explore and demonstrate the key benefits of 5G technology to transform the various vertical industries and enable innovative applications which will ultimately contribute to European Union digital transformation. 5G PPP has set many dedicated projects to engage with different vertical industries, capture their use case requirements, design a 5G based solution and validate it both from technology and business perspective in the form of prototypes or advanced trials. For a complete list of projects funded, see [i.32], [i.33] and [i.34].

To that end the 5G PPP launched three research infrastructure projects, namely 5G-EVE [i.35], 5GENESIS [i.36] and 5G-VINNI [i.37]. Each of them provides an end-to-end testing platform for vertical industries to validate a wide variety of use cases in both controlled and large scale setups. The platform's capabilities do not just account the different 5G network standard features to experiment with but also tools and processes to carry out their testing and measurement activities. These platforms help on processing the KPI requirements of verticals for deriving and automatically synthesizing and launching several test cases over their 5G facilities. The data generated by the execution of all those relevant test cases are gathered, analysed, and summarized for the vertical users to help them characterize the behaviour of their 5G-compatible applications and end-user devices, under a variety of internal and external conditions considered. A detailed overview of these platforms and a comparative assessment is provided in annex A.

These projects provide feedback to other relevant initiative, the 5G PPP Test Monitoring and Validation Work Group (5G PPP TMV WG), that has delivered a set of highly valuable recommendations on automatic testing framework for vertical KPI validation (including Testing as a Service (TaaS) approach), model driven methodology, and the mapping of E2E vertical service KPIs vs technical network KPIs. Clause 4.2.3.2 provides more details.

The Test, Measurement, and KPIs Validation (TMV) Working Group was founded as part of the 5G PPP effort to promote commonalities across projects that have strong interest in the T&M methodologies needed to provide support to the vertical use cases in the 5G Trial Networks. Such efforts include the development of Test and Measurement methods, test cases, procedures and KPI formalization and validation to the greatest possible extent, ensuring a unique European vision on how to support the entire lifecycle of the 5G network, from R&D to actual deployed environments.

The Group considers the following research areas and technology domains:

- Testing KPI definition, KPI sources, collection procedures and analysis.
- Testing frameworks (requirements, environment, scenarios, expectations, limitation) and tools.

- Testing methodologies and procedures.
- KPI validation methodologies.
- Testing lifecycle (i.e. testing execution, monitoring, evaluation, and reporting).
- Common information models for 5G Test and Measurement (T&M).

Another important topic is the use of and contribution towards open source projects such as Open Source MANO (OSM) [i.38], Open NFV(OPNFV) [i.39] or Open Network Automation Platform (ONAP) [i.40] as well as the identification of relevant exploitation and dissemination targets to promote a global adoption of 5G T&M.

4.2.3.2 5G PPP TMV

4.2.3.2.1 Testing Methodologies and Testing as a Service

TMV WG makes a fundamental discrimination between testing and monitoring. Testing provides a greater observability due to the active control over the type and intensity of traffic that is pushed through the network and through subsets of the network elements. This provides more degrees of freedoms in selecting what can be tested and measured (e.g. scalability or security resilience). Monitoring is instead a generally passive process that is providing metrics from various components/layers of the 5G network. For this reason, the KPIs that can be measured via testing are substantially different than through monitoring alone as described in [i.1].

TMV identifies the 5G network as a system composed by several complex and heterogeneous components, blending IT, cloud, and telecommunication technologies, stacked on top of each other to create the full 5G network like the pyramid in Figure 2. At the bottom of the pyramid there are the basic transport technologies, such as front- and backhaul, and the Data Center network fabric. On those, the NFV infrastructure is built, with cloud technologies such as OpenStack. The Management and Orchestration (MANO) is instead a kernel component for enabling the NFV principles. The telecommunication and service components seen as Virtual Network Functions (VNFs) can be then included in the picture, and with those, at the tip of the pyramid, there are the E2E network services.

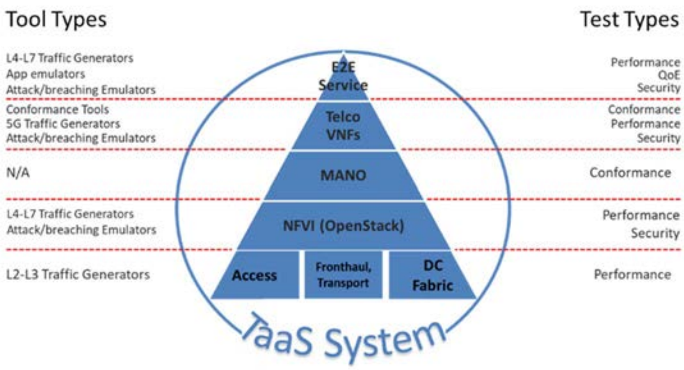

Figure 2: TaaS system overview

Each level carries along different types of tests that **should** be performed while deploying and integrating the network, onboarding the VNFs, and providing the services. In this direction, Test as a Service (TaaS) plays then an important role in reducing the effort that the MNOs' (Mobile Network Operator) engineers need to put in testing the 5G infrastructure and components. By simplifying the testing operations and providing an interface to connect to the Continuous Integration / Continuous Deployment (CI/CD) pipelines of the MNOs NetOps, TaaS is promising a stable performance delivery while maintaining under control (or even reducing) the OPEX. TaaS is expected to become an essential component of the Zero-Touch philosophy that is currently pursued in standards like ETSI ZSM [i.41], ETSI GANA (Generic Autonomic Networking Architecture(ETSI TS 103 195-2 [i.20])) related Autonomic/Autonomous Networking (ANs) standard(including GANA instantiations onto various network architectures and their associated management and control architectures), ETSI TR 103 747 [i.62], ETSI TR 103 473 [i.64] (V1.1.2), ETSI TR 103 404 [i.65],other standards on ANs outside of ETSI (e.g. Recommendation ITU-T Y.3324 [i.63], NGMN 5G End-to-End Architecture Framework v3.0.8 [i.49]), or open source communities such as OPNFV [i.39] and ONAP [i.40]. In Figure 2 it is also illustrated how a TaaS automation system can bind together different types of Test Tools, and which types of testing are covered. It possible to note that all the aspects ranging from Conformance to Security, from Performance to QoE can be test through a TaaS system, making it a powerful tool in the end of network equipment vendors, MNOs, and vertical customers.

The TMV recommends supporting and using a TaaS approach in the future 5G networks, and the group is currently working on identifying several commonly shared, standardized Test Cases useful at the validation of the E2E services.

Since most of the testing methodologies for the individual components are expected to come from the relevant standards, the TMV's priority is to provide methodologies and Test Cases for the validation of the E2E services delivered to the verticals. This is the ultimate goal of the present document which leverages and extends the TMV guidelines in order to provide recommendations on the functionality to be implemented by a generic vertical application validation framework in 5G and beyond networks.

4.2.3.2.2 Essential KPIs for Service Validation

The first step for creating the needed Test Cases in a TaaS framework for 5G service validation is to identify which KPIs should be stressed by the tests. In this direction, TMV WG identified a list of basic 5G technical KPIs, and for each of them TMV defined its type (monitoring/testing) and the related observing points in the 5G network. The outcome of this work is reported in the TMV whitepaper [i.1].

Based in that previous analysis, the TMV WG recently has released an exhaustive analysis of 5G PPP projects' use cases of various verticals mapping their service performance KPIs to corresponding 5G network KPIs. The use cases cover a wide number of vertical service areas, such as smart cities and utilities, transportation, automotive, media and entertainment, agriculture and agri-food, smart (air)ports, energy, and e-health and wellness. Such comprehensive analysis can be found in [i.42].

4.2.3.3 5G PPP validation platform solutions: 5G-EVE, 5GENESIS and 5G-VINNI

The 5G EVE platform offers an integrated set of tools, automated procedures and site facilities to allow Vertical industries to run their experiments in a 5G enabled infrastructure distributed over various site facilities, providing virtual testing environments easy to customize in terms of contexts, test cases, metrics and KPIs to be collected, etc., where vertical services can be validated in realistic scenarios.

The 5G-VINNI Facility consists of multiple, inter-connected sites, each of which supports demonstration of a range of KPIs, using specific access technologies and end-user equipment types. 5G-VINNI offers a testing infrastructure that is able to verify and validate the performance of the 5G-VINNI facility in terms of the 5G PPP KPIs. The testing infrastructure allows vertical customers to use the facility with a Testing-as-a-Service (TaaS) model, enabling the execution of dedicated campaigns with reduced effort. Open APIs and SDKs enable customers to integrate their own technologies within the framework.

The 5GENESIS platform implements and verifies evolutions of the 5G standard via iterative integration and testing procedures. Heterogeneous physical and virtual network elements are unified under a common coordination and openness framework that is exposed to experimenters/vertical industries and enables end-to-end slicing and experiment automation.

An overview of the three different solutions can be found in annex A.

4.2.3.4 Future Internet Research and Experimentation (FIRE)

The aim of both FIRE and FIRE+ EC ICT programs was providing support for building infrastructures for the design and deployment of products, applications, and services on the Future Internet.

Such infrastructures were aimed to achieve the following goals:

- experimental capability at European level that covers a variety of networking technology areas and allows tens of experiments to be run on top of them each year;
- potential to experiment without the constraints of the physical location or access to a specific experimental facility; reduction of the time to experiment by allowing a larger set of experiments to take place on reliable and benchmarked infrastructure that can evolve and be re-configured;
- response to the needs of individual, small and medium experimenters without access to experimental facilities or environments;
- support of trials driven by vertical application areas with a good mix of supply and demand stakeholders;
- contribution to the sustainability model of experimental facilities;

- contribution to standardization and interoperability of experimental facilities.

The project identified reference deployment scenarios, defined new KPIs (Key Performance Indicators) and QoE metrics, developed new testing methodologies and tools, and designed a complete evaluation scheme. The framework, methods and tools developed during the project focused on providing the mechanisms to incorporate new wireless technologies and topologies envisaged in 5G and contribute to the new ecosystem.

An overview of FIRE and FIRE+ can be found in Annex A.

4.2.4 Assessment of the state of the art solutions

4.2.4.1 Comparison of 5G PPP validation solutions

Table 1 summarizes the comparison of the experimental projects referred from several perspectives.

Table 1: Comparative table of state of the art solutions

	5G EVE	5G VINNI	5GENESIS	5GinFIRE
Overall objective	5G EVE provides a multi-site 5G validation platform to evaluate the performance of end-to-end vertical services in flexible 5G environments. 5G EVE offers a wide set of vertical-oriented functionalities to simplify the evaluation of service and network KPIs, including intent-based interfaces for experiment definition, automation of test execution and open APIs to integrate new facilities and orchestration platforms.	To build an open large scale 5G End-to-End facility that can demonstrate that key 5G network KPIs can be met, and be validated, accessed and used by vertical industries to test use cases and validate 5G KPIs.	5GENESIS project has specified and implemented an experimentation methodology focused on the validation of 5G Key Performance Indicators and the quantification of the performance improvements introduced in the verticals solutions under test.	FIRE projects provided infrastructures for the design and deployment of products, applications, and services on the Future Internet.
Validation framework	Testing-as-a-Service (TaaS) platform, supporting custom experiments, configurable virtual services and network slicing, collection and evaluation of service and network KPIs, integrated diagnostics.	Testing-as-a-Service (TaaS) platform.	Testing-as-a-Service platform supports the execution of standards experiments defined by the project and custom experiment defined based on specific testing requirements coming from the verticals. Scenarios and network slicing configurations are also provided but can be defined new ones. The experiments can be fully automated or can include the human intervention of the final user of the services under test. Finally, a monitoring and analytic module is included to provide advanced analysis of the measurements collected.	FIRE projects provided infrastructures for the design and deployment of products, applications, and services on the Future Internet.

	5G EVE	5G VINNI	5GENESIS	5GinFIRE
5G enabled testing environment	Multi-site facility distributed in Spain, Italy, France and Italy, managed through an interworking layer to unify the management of inter-site and intra-site experiments. 5G EVE facilities offer the deployment of the 5G architecture, with different spectrum and access technologies, support for network slicing, edge computing, NFV orchestration and network monitoring. 5G EVE infrastructure can be extended with additional sites through its interworking layer.	Facility sites in seven European countries. Testing can be performed within a Facility site or be performed as inter-site tests.	The 5GENESIS experimentation framework is devoted to the full control of a facility but enabled the definition of distributed experiment between different facilities.	FIRE projects were not specifically focusing on 5G but to Future Internet in general. For example, Fed4FIRE+ is offering the largest federation worldwide of Next Generation Internet (NGI) testbeds, supporting a wide variety of different research and innovation communities and initiatives in Europe, including the 5G PPP projects and initiatives. TRIANGLE is building a framework to help app developers and device manufacturers in the evolving 5G sector to test and benchmark new mobile applications in Europe utilizing existing and extended FIRE testbeds.
Experiment workflow	• Experiment design • Experiment preparation • Experiment execution and monitoring • Experiment results evaluation	• Experiment Design • Experiment preparation • Experiment execution • Experiment assessment	• Experiment consultation phase • Experiment provisioning phase • Experiment execution – Pre-run – Run – Post-run • Experiment decommissioning phase • Analysis of the results	Fed4FIRE+ Experiment lifecycle management as a service: • Resource discovery • Resource specification • Resource reservation • Resource provisioning • Experiment control • Monitoring • Measuring • Permanent storage • Resource release

	5G EVE	5G VINNI	5GENESIS	5GinFIRE
Testing automation capabilities	5G EVE Portal provides a single access point to request the scheduling of vertical experiments, trigger the automated deployment of the virtual environments and launch the automated execution of the experiment test cases, which includes the collection and validation of service and network KPIs and diagnostic analysis.	Tests can be scheduled in the TaaS user interface. TaaS can also be programmatically controlled through REST APIs.	The 5GENESIS experimentation framework includes an entity called Experiment Lifecycle manager which enables automating the execution of the experiments.	Fed4FIRE+ provides a set of tools to enable easy configuration and execution of experimental set-ups on a wide range of Fed4FIRE+ testbeds. Fed4FIRE+ testbeds can be fully operated remotely, where the only technical requirement for experimenters is to have standard Internet connectivity. TRIANGLE offers a direct access to the Keysight TAP (Testing Automation Platform), which is a programmable sequencer of actions with plugins that expose the configuration and control of the instruments and tools integrated into the testbed.
KPI collection and validation	Support for collection of both service and network KPIs (user data rate in DL/UL, RTT latency, reliability and availability), which can be visualized through graphs in the 5G EVE portal, collected in real-time via REST APIs or downloaded for further processing. The platform integrates mechanisms for threshold-based KPI validation and performance diagnostics.	Measurement results available through a Grafana application or stored to external databases. Thresholds can be set for KPI validation.	The analytics component of the Coordination Layer is responsible for the complete collection and analysis of the heterogeneous monitoring data produced during the usage of the 5GENESIS experimentation. In order to collect the monitoring information from all the elements of each 5GENESIS platform, the analytics component retrieves the measurements from the probes deployed in each platform. This component ingests either in-real time or after the end of each experiment session, the measurements in a unified database for post-processing and long term storage. Raw results are available to be downloaded. The analytics module provides a powerful correlation tool developed to provide advanced analysis.	In Fed4FIRE+ semantic resource directory is used to collect, transform and offer monitoring information about resources. It could be also used for storing measurements coming from probes deployed by the experimenters. In TRIANGLE a testbed management framework provides full test case automation, by coordinating testbed component configuration, their execution, processing the measurements made in each test case, and computing QoE scores for the application tested.

	5G EVE	5G VINNI	5GENESIS	5GinFIRE
Experiment definition	Experiments are defined through configurable blueprints, which define vertical service elements and connections, service parameters, network requirements, context conditions, service and network KPIs, target values and evaluation criteria, test cases and configurable test scripts. Experiments can integrate custom VNFs/PNFs.	• Purpose • Description • Initial Conditions • Parameters • Procedures & Expected Results	An experiment descriptor template has been elaborated to allow an univocal definition of the experiment. The experiment descriptor is orchestrated around three key concepts: test cases, scenarios and slices. The information needed to fill the template is retrieved from the date introduced by the vertical in the 5GENESIS Portal. The experiment descriptor can be also provided directly via the open APIs.	Fed4FIRE+ provides a set of tools to enable easy configuration and execution of experimental set-ups on a wide range of Fed4FIRE+ testbeds.
Interfaces to Verticals and Experimenters	5G EVE offers a web-based Portal for the design, scheduling, deployment, execution and verification of 5G experiments. The Portal provides both an Intent-based interface and a wizard for the experiment definition. REST APIs are available for the management of experiments from third party systems. KPIs can be visualized on the web portal, downloaded or accessed via REST API.	TaaS offers both a GUI and a REST API for composing and scheduling tests. Results can be presented in tools offered by 5G-VINNI or stored to external databases.	The Experimenter can set experiments and get results through the 5GENESIS Portal, as well as directly execute the experiments via the open APIs.	Fed4FIRE+ provides a set of tools to enable easy configuration and execution of experimental set-ups on a wide range of Fed4FIRE+ testbeds.

4.2.4.2 Top-5 key features in state-of-the-art validation platforms

Table 2 highlights the top-5 more relevant features contributed by 5G EVE, 5G-VINNI, 5GENESIS and FIRE as validation platforms.

Table 2

Platform	Top-5 key features
5G-EVE	• Multi-site validation platform with facilities distributed in Spain, Italy, France and Greece providing full deployment of 5G networks, easy to extend with additional sites through a plugin-based interworking layer. • Vertical-driven approach for experiment definition, with intent-based interfaces and wizards to build service and experiment blueprints. • Fully automated management of the entire experiment lifecycle, including service deployment, experiment configuration and execution, service and network KPI collection, results evaluation and analysis. • Complete set of REST APIs for service and experiment deployment, configuration, execution and monitoring, to enable the integration with 3rd party platforms and orchestration systems with support for closed-loop interactions and automation. • Rich set of tools for service validation, including experiment reporting, configurable KPI assessment, support for external KPI processing and integrated diagnostics and analysis.

Platform	Top-5 key features
5G-VINNI	• Facility sites available for testing in seven European countries. • Offers a trial end to end facility of the latest 5G technologies for radio. • Access (including mm-wave), backhaul and core networks (both NSA and SA). • Leverages advanced virtualisation technologies and optimization algorithms. • Offers a comprehensive set of tools both for network and application level testing. • Verticals' own testing tools can be integrated into the 5G-VINNI testing-as-a-service platform.
5GENESIS	• Facility sites available for testing in five European countries and a common experimentation framework is offered on top off all of them. • The experimentation framework has been defined to facilitate the execution of the experiments by verticals, abstracting the complexity of the low layers of the experimentation framework. The verticals can set experiments and get results through the 5GENESIS Portal, as well as directly execute the experiments via the open APIs. Defining the experiments via the Portal is probably the preferred option for most of the Experimenters, as they will be able to run experiments in a controlled environment all the times they want, and, if needed, automate the execution via the open API. • 5GENESIS experimentation framework support the testing over heterogeneous networks and the testing of heterogeneous verticals solutions. • 5GENESIS experimentation framework is driven by the execution of test cases and addresses the design and implementation of procedures for executing properly defined experiments and test cases which are based on the vertical requirements. The 5GENESIS experimentation framework has a modular architecture that can be easily reused in other 5G testbeds also thanks to the open-source nature of their software components which are part of the of the Open 5GENESIS Suite available at https://github.com/5genesis.
FIRE	• FIRE provides access to external users to test their own applications. • FIRE federates several local testbed. • The scope covers Next Generation Internet (NGI), in general. • FIRE provides contemporary access to several experimenters by providing them with resource slices. • FIRE offers experimenter support.

4.2.4.3 Top-5 potential enhancements in state-of-the-art validation platforms

Table 3 remarks the top-5 more important enhancements contributed by 5G EVE, 5G-VINNI and 5GENESIS as validation platforms.

Table 3

Platform	Top-5 enhancements
5G-EVE	• Enhanced flexibility in the programmable management of experiments and service lifecycle (e.g. support for scaling actions, network and service function re-configurability, etc.). • Extended support of network slicing and dynamic radio network configurations. • Support of network programmability and orchestration features exposed to third party systems. • Enhanced configurability of test cases, experiment configuration and procedures. • Reduced complexity of experiment design.
5G-VINNI	• Support for advanced RAN features such as slicing, URLLC and positioning (subject to availability). • Extend the set of testing tools available through the testing-as-a-service platform. • Enhance automation in the testing-as-a-service platform. • Enhance mm-wave support with more stable mm-wave devices. • Integration with Non Public Networks (NPNs).
5GENESIS	• End-to-end automation of the experiment workflow. • Open sourced experiment coordination tools, slice manager and performance monitoring tools. • Well-defined control plane interfaces to enable the expansion of the platform with new components. • Support for the deployment of vertical services. • Portal for verticals where the experiments can be defined without dealing with low level configuration details.
FIRE	• They provide access to external users to test their own applications. • They are federating several local testbed. • Their scope covers Next Generation Internet (NGI), in general. • They provide contemporary access to several experimenters by providing them with resource slices. • They offer experimenter support.

4.2.4.4 Assessment conclusions

The 5G validation platforms developed by the three 5G PPP platform projects together with the FIRE projects present major common aspects in terms of experimentation methodology, framework design and infrastructure capabilities, with each of them proposing specific added value features that extends the basic set of functionalities on the basis of specific requirements from their target vertical services and use cases.

In general, all the validation platforms are offering a testing infrastructure distributed on multiple sites across Europe, each of them providing a complete 5G network equipped with the latest 5G technologies and a wide set of orchestration and monitoring tools. Some infrastructures may deploy site-specific technologies for advanced testing in particular context conditions (e.g. mmWave or satellite networks) and/or offer the possibility to extend the original testbeds with additional facilities (e.g. supporting the integration with Non Public Networks or through a modular interworking layer that can interconnect to external sites).

The validation frameworks are based on the common concept of "Testing as a Service", supporting different levels of customization for the definition of the experiments. In most cases, it is possible to define the service components, the target environments and the characteristics of the network slices where the service needs to run, the KPIs to be collected and test cases to be executed. Similarly, all the platforms provide a certain degree of automation in the different steps of the experiment lifecycle. In particular, the basic service provisioning and infrastructure configuration supported by all the platforms, is integrated with automated procedures for configuration and execution of test cases, together with the automated collection of KPIs. In some cases, the KPI validation is also automated and enriched with added-value tools for performance diagnostics.

Another key aspect of these projects is the type of interface they offer to potential customers, i.e. verticals and experimenters. Here the key is to provide the right compromise between the simplicity of the experiment request (particularly important for the verticals) and the flexibility of the experiment customization options. In this sense, the platforms are offering custom interfaces that try to mix several tools and methodologies to approach the experiment management, e.g. mediating the user interaction with the platform through web-based portals, or offering intent-based interfaces to define the experiments. On the other hand, the offer of REST APIs to enable the programmable interaction with 3rd party systems is the preferred choice to support more complex experimentation actions driven by external platforms.

The following clause elaborates a number of recommendations based on the experience of 5G-EVE, 5G-VINNI and 5GENESIS projects.

4.3 Recommendations for the Validation Framework

4.3.1 Introduction

The main system proposed for meeting the expectations of the relevant Stakeholders of the validation of vertical applications over 5G and Beyond networks introduced in clause 4.1 (in a generic manner it will be referred as B5G application validation framework, B5G-app VF). The B5G-app VF is therefore a composite system of components, services and tools operated by CSPs and meant to be used by partner/engaged Verticals. A reference model and recommendations for building out, operating and using such a validation framework are described in next clause (clause 4.3.2).

The following recommendations are provided on the base of the current state-of-the-art. The evolution to B5G solutions could bring novelties that could require updates of the present document.

The following sub-sections of the present document outline the deployment and functional reference model of the Validation Framework (VF) and recommendations on segmented aspects of Validation Frameworks (VFs) over B5G. Those aspects are:

- Deployment and reference model.
- Validation framework capabilities.
- Validation Processes.
- KPI mechanisms.
- Vertical applications design.

4.3.2 Recommendations

4.3.2.1 Deployment and reference model

The B5G-app Validation Framework can be deployed in a B5G trial environment such as the case of 5G PPP programs or in a B5G commercial network. In both cases the network infrastructure will account a MANO orchestration system. In case the site facility does not provide a CSP or third party application service hosting environment, the B5G-app VF can host the vertical application.

Recommendation B5G-app VF Deployment #01: Principles of the reference model for the experimental platform

It is recommended that the B5G-app Validation Framework should follow the below solution principles:

- Provide a broad suit of services specifically designed for validating applications leveraging B5G networks, as introduced in the Motivation in clause 4.1 and depicted in Figure 1.
- Enable an end-to-end vertical validation process, from the inputs and requirements of the vertical to the delivery of validation results (Portal and APIs toward vertical experimenters/testers).
- Provide an open environment to the Vertical, meaning that vertical application systems can be deployed for being validated, using open interfaces and APIs (Portal and APIs toward vertical experimenters/testers).
- Deliver experimental results of general validity and applicability in B5G networks, obtained through a set of uniform process and mechanisms which are independent of the specific technology vendors involved in its underlying infrastructure.
- Offer validation services of general application to all kinds of vertical applications, without making any assumptions on the specific nature of the vertical services to be validated.
- Adopt and leverage relevant standards (e.g. MANO related standards) at both infrastructure and service levels.
- Provide access to a well-defined set of B5G exposed capabilities that make it possible to carry out validation activities over full-fledged advanced and end-to-end network setups.
- Abstract the complexity of the B5G underlying infrastructure and the validation process implementation by using open solutions for control and management (e.g. MANO) by offering easy to use interfaces to the Vertical experimenters (Portal and APIs toward vertical experimenters/testers).
- Expose capabilities and tools for enabling a wide range of validation conditions and strategies to the Vertical experimenter (by utilizing the capabilities of e.g. an experiment configuration component).
- Provide automation capabilities for recurrent and otherwise error-prone tasks involved in validation activities (using the functionalities of e.g. experiment execution, and experiment configuration analysis/validation components).

Figure 3 illustrates the B5G-app Validation Framework Reference Model including main functional blocks.

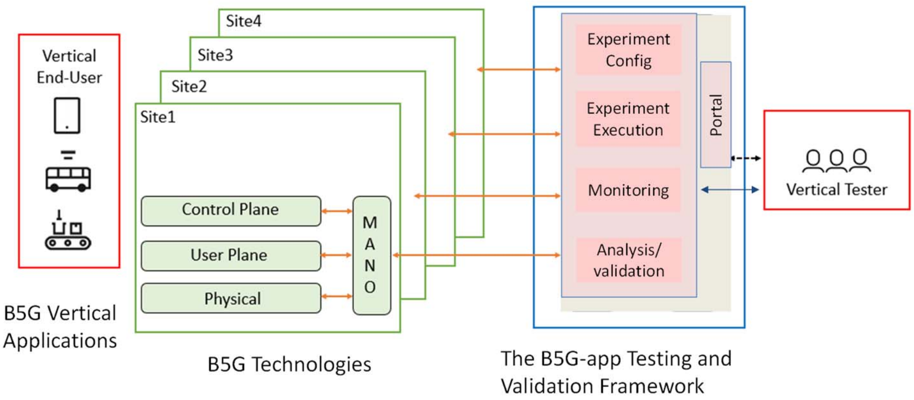

Figure 3: B5G-app VF Reference Model

4.3.2.2 Capabilities

4.3.2.2.1 5G Capabilities & Enablers

Recommendation B5G-app VF Capabilities #01: 5G NR capabilities

A B5G-app validation framework should enable the verticals to test their services under different 5G NR deployments and take fully advantages of the main features introduced by 5G NR:

- New Bands and increased bandwidths: NR can be deployed in a very range of bands both in existing IMT delivered intervals and in future bands. The differences between bands are very pronounced for NR due to the very wide range of frequency bands.

  Frequency bands within the scope of the present Release 15 work in 3GPP are divided into two frequency ranges:

  - Frequency range 1 (FR1) includes all existing and new bands below 6 GHz.
  - Frequency range 2 (FR2) includes new bands in the range 24,25 - 52,6 GHz.

NOTE: Refer to [i.43] and [i.44] for further information on 5G NR radio transmission and reception bands.

- Massive MIMO & beamforming: Multiple-input and multiple-output (MIMO) is a key technology to improve throughput. It uses multiple antenna arrays both on the transmitter and on the receiver sides, as to enable multi-layer data transmission.

  NR supports multi-layer data transmission for a single UE (single-user MIMO) with a maximum of eight transmission layers for DL and four ones for UL. NR supports also multi-layer data transmission with multiple UEs on different layers (multi-user MIMO) with a maximum of twelve transmission layers for DL and UL transmission.

  Since NR supports multi-beam operation where every signal/channel is transmitted on a directional beam, beamforming is an important technique for achieving higher throughput and coverage especially, in high-frequency range.

- Multi-Services transmission: A very wide range of deployment scenarios has been considered for 5G; from large cells with sub 1 GHz carrier frequency up to mm-wave deployments with a large spectrum allocation. A flexible OFDM numerology (μ) with subcarrier spacing ranging from 15 kHz (used in LTE) up to 240 kHz has been considered in ETSI TS 138 211 [i.45]. Different numerologies can be used simultaneously in a cell. Compared to LTE, higher carrier spacing allows achieve lower latency in the air interface.

  Despite such slices are running on the same physical network from the end-user point of view they appear as independent networks and each of them may provide different network capabilities.

  The characteristics of each slice are defined in terms of QoS, bit rate, latency, etc. For a given slice, these characteristics are either predefined in the 3GPP Standard or are operator-defined.

  There are three types of predefined slices: type 1 - is dedicated to the support of eMBB, type 2 - is for URLLC and type 3 - is for MIoT support. These predefinitions allow inter-PLMN operation with reduced coordination effort between operators. As for the operator-defined slices, they enable more differentiation among Network Slices. A dedicated Network Function in 5G Core Network is introduced for handling slices: the "Network Slice Selection Function" (NSSF), which enables the selection of the appropriate slice(s). The UEs may use multiple Network Slices simultaneously.

Recommendation B5G-app VF Capabilities #02: Network Slicing capabilities

One of the 5G Key technology enablers which has to be considered as a Service Capability for a B5G-app validation framework should be Network Slicing. Network Slicing can be described as a mean *"to satisfy the demand of dedicated telco services with specific Service Level Agreements (SLA)"*. It is a way to ensure use case performance requirements described by the Vertical.

The service could be described as "Network Slice as-a-service" which provides a concrete answer to Vertical's demand by enabling *"à la carte"* End-to-End services. In this model, a Network Slice is offered by a Communication Service Provider (CSP) to a Communication Service Customer (CSC) for a communication service that is based on a Network Slice Instance. In B5G-app validation framework, the manager of the B5G facility would be the CSP and the Vertical/Experimenter would be the CSC.

The concept of Network Slicing has been defined as a key feature for 5G by 3GPP in [i.46]. It has been also defined by Recommendation ITU-T Y.3100 [i.47] and ETSI TS 123 501 [i.48] as "*a logical network that provides specific network capabilities and network characteristics*" with following interesting notes for ITU:

"*NOTE 1 - Network slices enable the creation of customized networks to provide flexible solutions for different market scenarios which have diverse requirements, with respect to functionalities, performance and resource allocation.*

NOTE 2 - A network slice may have the ability to expose its capabilities.

NOTE 3 - The behavior of a network slice is realized via network slice instance(s) (NSI)."

In this context, a Network Slice Instance (NSI) is a set of network functions and the resources for these network functions which are arranged and configured, forming a complete logical network to meet certain network characteristics.

In this direction, Network Slice "logical network" within a B5G-app validation framework should include 5G system network functions (consisting of 5G Access Network (AN), 5G Core Network and UE), but also additional network functions needed to fulfil the SLA of the service within this logical network. In cases where a B5G-app validation framework will operate end-to-end Network Slices and 5G services across multiple administrative domains, the Network Slice "logical network" should include also the transport network that is used to interconnect 5G system network functions between different administrative domains.

Note that on the same physical core and radio networks, different slices can run as, for example, one supporting mobile broadband application in full mobility, as provided by the legacy LTE system, and another slice delivering as an example, non-mobile, latency-critical industry-automation application.

Recommendation B5G-app VF Capabilities #03: Cloud and Edge Computing capabilities

A B5G-app validation framework should include Cloud and Edge computing capabilities. Edge Computing is generally referred to as a distributed computing paradigm where computation is largely or completely performed on distributed device nodes as opposed to a centralized cloud environment: edge computing pushes applications, data and computing power (services) away from centralized points to the logical extremes (closer to end-user) of a network.

The testing of vertical service in a B5G-app validation framework which includes Cloud and Edge computing capabilities will have the following benefits:

- Latency reduction, that is the time needed by data to travel from source device to the place where they are elaborated (today, generally the "cloud").
- Traffic volume reduction, as the local elaboration of big amounts of data, may provide a significant reduction of data moved over the network. By deploying various services and caching content at the network edge, core networks are alleviated of further congestion and can efficiently serve local purposes.

These benefits make Edge Computing to be an enabler or at least, a key feature to achieve performance requirements for some use cases envisioned in 5G, where very low latencies are required and/or the high increase expected of simultaneous devices sending / receiving data, imposing stringent high capacity requirements in the transport network.

3GPP has identified this in System Architecture specification ETSI TS 123 501 [i.48], by addressing different mechanisms to support Edge Computing. As described in the present document, Edge computing enables operator and 3rd party services to be hosted close to the UE's access point of attachment and so, to achieve an efficient service delivery through the reduced end-to-end latency and load on the transport network. This can be achieved by implementing Local Break Out at the Edge infrastructure to connect the data plane to the Applications.

The testing of vertical service in a B5G-app validation framework which includes Cloud and Edge computing capabilities will have also the benefit that applications will be deployed in an environment very much like the modern cloud the developers are familiar with; ideally, they should not face great differences working with a commercial cloud environment and an "Edge" cloud. In this sense, the enabling key is the exposure of services by mean of RESTful (REST) APIs and the availability of virtualization and/or containerization technologies. Modern mobile networks are already familiar with these concepts, in fact, network deployments are today focusing on NFV technology and different sets of APIs.

ETSI has created a set of standards to define Multi access Edge Computing solutions [i.49], and there are several industry initiatives to provide software stacks based in Virtualization technologies (VMs, Containers) such as Akraino [i.50] or Openness [i.51].

Recommendation B5G-app VF Capabilities #04: NFV&SDN Capabilities

A B5G-app validation framework should include SDN and NFV capabilities. Network Function Virtualisation (NFV) and Software Defined Networks (SDN) are two pillars of 5G and Beyond networks to target to meet the requirements of a highly mobile and fully connected society. Both technologies, SDN and NFV, address fundamental 5G demands concerned with high network flexibility as well as a service-driven approach. NFV was developed by service providers with a goal of accelerating an introduction of new services on the networks. Proprietary equipment applied in traditional networks by CSPs made impossible quick provisioning of new services. SDN has grown from an approach of programmable networks. In a programmable network, a behaviour of network hardware and flow control is handled by software that operates independently from network hardware. A goal of these networks is to enable re-programming (using well-defined APIs) a network infrastructure instead of having to re-build it manually.

NFV and SDN are mutually beneficial but are not dependent on each other. On one hand, Network Functions can be virtualized and deployed without an SDN being required, and on the other hand the network can be programmable connecting non-virtualized Network Functions. NFV benefits from SDN role in orchestration NFVI resources through features such as provisioning and configuration of network connectivity and bandwidth, automation of operations, security and policy control, with SDN being a key enabler technology of NFV Virtual Infrastructure Manager (VIM) function.. SDN benefits from NFV-introduced concepts such as virtualized infrastructure managers and the orchestrator given that an SDN controller could run on a VM.

In the B5G-app validation framework infrastructure SDN-based network control should enable dynamic programming of physical networks in transport as well as radio domains, based on one or more SDN controllers.

ETSI NFV [i.52] and Open Source projects such as OSM [i.38] or ONAP [i.40] define different solutions to orchestrate and automate the deployment of cloudified 5G infrastructure and standards interfaces for the B5G-app validation framework infrastructure to integrate with.

Recommendation B5G-app VF Capabilities #05: Interconnection and Extension capabilities

A B5G-app validation framework should include Interconnection and Extension capabilities of B5G facility that is composed of B5G infrastructure and the vertical services. Indeed, a secured connectivity should be put in place between the "B5G testing and validation platform" to the B5G infrastructure platform where the service is tested and validated. Typically, network connectivity ensures the network availability between the north bound interface (connected to the B5G testing and validation platform) to the orchestrator managing the B5G facility. In order to be able to interconnect several B5G Facilities infrastructures to the validation platform, it is convenient to adopt some specific prerequisites such as:

- To implement L2/L3 secured VPN. Manage a specific process for configuring the network equipment (routers, firewalls) and key security exchange. This part is very dependant to the type of equipment that are used at each facility as well as the rule security policies of each site.
- To define the main protocols, ports, IP addresses to be used.
- To define the control and data planes subnets when service testing is carried out in several B5G facilities.
- To be able to monitor some probes put at the network infrastructure and/or service levels in order to provide, in real time, the KPIs metrics back to the B5G testing and validation platform.
- To be able to provide to the B5G testing and validation platform one way (virtual machine for instance) to supervise the network connectivity, infrastructure up and running, by making some connectivity tests (data throughput, latency and jitter) via some probes and/or tools (iPerf tests [i.53] for example).

- To be able to forward KPIs performance to the portal to provide to the experimenter the performance values measured during the experience.

These recommendations should allow to extend the B5G facility cluster to several infrastructures.

Because the different B5G facilities are not implementing the same orchestrator, it is crucial to adopt a common interface for the deployment of virtualized components from the B5G testing and validation platform to the B5G facility. Typically, the ETSI NFV-SOL 005 [i.54] interface is used for managing the multi-orchestration part.

Recommendation B5G-app VF Capabilities #06: B5G Exposure Capabilities

B5G-app VF is recommended to take advantage of all the available observation points from the core and edge part of 5G and Beyond infrastructures. This recommendation refers to the capability the B5G-app VF to request and extract vertical-interested experimentation data through APIs exposed from the mobile core and the edge network.

The set of exposed APIs to be consumed by the B5G-app VF are vertical- and experiment- specific.

Focusing on the mobile core, the standardized northbound APIs that are provided by ETSI TS 129 522 [i.55] or other core network functions are to be consumed by the B5G-app VF. The capability for a standardized data exchange between the B5G-app VF and the 5G core of the B5G infrastructure is prerequisite to the APIs consumption. Thus, the entity of the B5G-app VF that is to consume those APIs, is recommended to abide by standardized interaction frameworks set by the infrastructure side, such as for instance the (e)CAPIF framework ETSI TS 129 222 [i.56]. 3GPP SA6 is working on 3GPP TR 23.758 [i.57] to add in 3GPP Release 17 the architecture to support Edge Applications over 3GPP networks using CAPIF for API exposure.

Focusing on the edge part of the network, the utilization of vertical-interested experimentation data through APIs exposed from the 5G RAN can expands the measurements observation potential of the B5G-app VF. In this view, the exploitation of the ETSI MEC APIs provides the B5G-app VF with data and information that refer to the (radio) access part of the network, i.e. with information from the part of the network that highly impacts the end-to-end performance. Relevant MEC APIs are for instance ETSI GS MEC 012 [i.58] and ETSI GS MEC-DEC 032-3 [i.59].

Overall, the capability to use 5G core and edge interfaces as observation points for the B5G-app VF, guarantees:

- The use of the B5G-app VF with commercial standardized infrastructures and with facilities, where no direct access, further to the standardized interaction, is granted.
- The enrichment of the measurement campaigns conducted via the B5G-app VF with network level information i.e. with performance KPIs gathered from standardized APIs exposed by the infrastructure.
- An analysis on the achievable vertical-oriented performance results in relation to the network configurations and management choices, i.e. in relation to information available at the infrastructure side.

4.3.2.2.2 Testing and Validation Environment

The following recommendations refer to the testing and validation environment.

Recommendation B5G-app VF Capabilities #07: Application Deployment Environment

Each vertical application under test specifies the requirements for the needed environment and conditions in a blueprint document. In order to provide a proper execution and test environment to the specified application, a B5G-app validation framework could build and deploy a virtual environment supporting all the features specified in this blueprint document.

The virtual environment includes virtual machines, virtual networks and virtual functions, so it is important to define the needs of the application relating to these terms. For example, and considering what is already used for cloud applications, conditions such as CPU or GPU power, disk space, and connectivity should be met.

The requirement may include specific hardware capabilities, such as precise radio functions, or the presence of camera. The deployed environment should make sure that it includes the required equipment.

Moreover, the environment also needs to respect possible geographic constraints linked to the application and ensures that it starts on equipment that answer this requirement.

The application itself is defined as one or several functions or pieces of software made available in a series of binaries or scripts to be launched inside the virtual environment. Examples of ways to access to the application components include container images stored in private or public registries. Some components of the application might also need dependencies that are stored on public internet repositories. The framework should be able to retrieve the different parts of the application in order to launch them on the right target equipment.

Security is another point to consider as it is important to ensure that the access to the Application Environment is limited to the authorized users or networks.

Recommendation B5G-app VF Capabilities #08: Application Testing Environment and Tools

A B5G-app validation framework should provide means to control the different actions involved in the application testing. These actions include starting and stopping either the whole test or each different step of the test.

One possible way to achieve this is to provide an API (Application Programming Interface). This API could provide the functions for starting and stopping a scenario applied to a selected application.

The B5G-app validation framework might also provide a GUI (Graphical User Interface), using the API or not, that will allow a human operator to easily access the different actions of test.

For certain test cases, in order to create specific conditions in the virtual environment, it may be needed to simulate network load, for instance to see how the application will react. The deployment of virtual functions such as traffic generation, traffic control or traffic sink inside the environment should be made possible.

In order to understand what happens during the execution of the test, the B5G-app validation framework should include a way to log all the actions and their outcome, and also the possible encountered problems. The logging might be activated at different levels depending on the needed details, for instance: informational, warning, debug, etc.

Recommendation B5G-app VF Capabilities #9: 5G Network KPIs and Metrics Framework

Refer to clause 4.3.2.2.4.

Recommendation B5G-app VF Capabilities #10: Performance Diagnostic capabilities

Due to the increased commercial interest in 5G infrastructures there has also been an increase in the interest in software solutions to help deliver reliability in the components of the 5G network, as well as on the services running on top of the 5G network. Therefore, the research into the fields of performance diagnosis and root-cause analysis has been gaining popularity with hopes to find effective methods and models to provide reliability to the 5G services. The hope is that by predicting and localizing faults and service degradations, engineers and technicians can make fact-based decisions on how to improve the system or mitigate the possible faults. This in turn would allow for companies to deliver more reliable services. These features, which are critical for the deployment and delivery of 5G services, cannot be absent from a testing and validation platform.

However, understanding and predicting the performance of a service on the network and on the cloud is by its nature a hard thing to do. The services are often a part of a large and complex software system located in different virtual and physic entities across the 5G network. Therefore, understanding the performance of a system of that magnitude does not only require expert domain knowledge but also analytical models that often tend to be overly complex.

In order to address the aforementioned challenges, an advanced performance diagnosis mechanism based on enhanced data analysis, machine learning and AI should become part of the testing and validation environment for 5G and Beyond services. The developed diagnosis mechanism should offer insights regarding the observed performance and suggesting tips for performance improvement (e.g. by applying post-process analytics on the collected KPIs).

The performance diagnosis capabilities should include at least:

a) mechanisms for collection of additional metrics (in addition to the KPIs which are important for the vertical) in order to create critical relations between metrics and provide important insights;

b) mechanisms for anomaly detection and;

c) echanisms for Root Cause Analysis. A final, maybe optional, step is the automatic creation of optimizations actions which can be executed manually or automatically.

Additional metric collection

In addition to the metrics that are characterized as KPIs and are important to validate the performance (e.g. latency, throughput, availability, reliability), in a performance diagnosis approach additional metrics are important and are critical for the detection of an anomaly in the service or for identifying the cause of a problem. These metrics may include node metrics like CPU, MEM, Disk Usage and throughput measured on the interfaces of the network node, or latency metrics measured on the sub-paths of the actual service path.

Anomaly detection

Anomaly detection is an important data analysis task that detects anomalous or abnormal data from a given dataset. It is an interesting area of data mining research as it involves discovering enthralling and rare patterns in data. It has been widely studied in statistics and machine learning, and also synonymously termed as outlier detection, novelty detection, deviation detection and exception mining.

The first step of performing diagnosis in a 5G environment after collected all the required metrics (including the additional metrics) is the anomaly detection. Before the data can be analysed, and the process of finding the root-cause of a service degradation can be started, the system should first be able to detect that an actual anomaly is present.

In order to identify any anomalies that need to be considered further, the anomaly detection module should analyse the collected metrics of the experiments against a set of network profiling results and service profiling results which illustrate the normal execution of the service. The network/service profiling results should be generated a priori through the process of network and service profiling, and the results should be stored to be used for the anomaly detection mechanisms.

Root Cause Analysis (RCA)

The RCA mechanisms are responsible to predict and localize faults and service degradations, so that in a next step the engineers and technicians can take decisions on how to improve the system or mitigate the possible faults. The RCA module can utilize diverse information including correlated network/service events and E2E service graphs. By correlated network or applications events, it refers to events generated by different sources that can be related e.g. in temporal or spatial way. For example, events from neighbouring nodes or events from the same source and subsequent time slots. In addition, service graphs can be used as additional knowledge for the RCA algorithms in order to correlate nodes or link along a network path.

Optimization generations

The last step of the performance diagnosis mechanism is the generation of network or service optimization actions, which are in practise the steps and configurations that can be realized for the mitigation of possible faults and service degradations. During this step, the network and service profiling results can be used in order to decide the network configuration that will alleviate the issues or underperformance and return the network to a normal status. In a test and validation environment the outcome of this process may be the generation of a set of suggestion to the verticals. In a more advance system, this optimization actions may be enforced automatically by the system (e.g. through the network orchestrator).

Recommendation B5G-app VF Capabilities #11: Interfaces to Vertical Experimenters

An open API is the main interface for experimenters to define and execute their experiments. However, a Portal with a friendly Web User Interface (UI) can be provided to make the interaction with the Facility even easier. Such Portal should be able to display the execution logging output for all execution stages of the experiments (Pre-Run, Run and Post-Run). The foremost requirement of the open API is to present a common and open method of interaction between the 5G vertical or the experimenter - both commercial as well as experimental UEs - and the experimentation facility. It is recommended that such an API should provide the following operations:

- Access to the facility (including authentication).
- Experiments Definition.
- Experiments Configuration.
- Experiments Execution.
- Experiments Status monitoring.
- Experiments Results collection.

At least Portal should contain the following main blocks, as reflected in Figure 4:

- Register/Login: Registration and login pages for accessing the Portal.
- Main page: The main page shows all the experiments that the user has created, as well as the system notices and a log of the latest performed activities.
- Create Experiment: The experimenters can create their own experiments by configuring the available parameters. After that, the experiment can be run and examined, either by checking the status and execution logs or visualizing the results in the Grafana dashboard.
- VNF/NS Management: Users can upload their VNFs and NS to later use in the experiments. VNFs require an image and a VNFD (VNF Descriptor), while network services require an NSD file. This clause also allows users to remove their previously uploaded VNFs and NS. This functionality is not yet connected with the lower layers and appears as a proof of concept of the UI.

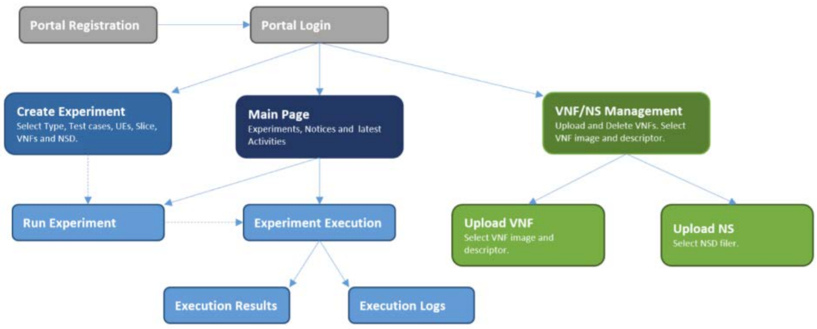

Figure 4: Portal architecture overview

The Portal should be able to display the execution's logs for all execution stages of the experiments. The user should be able to filter the messages displayed depending on the severity level. Those are: info, debug, warning, error and critical. Each line of the log should show the timestamp of when the message was produced, the severity level and the message itself.

For each experiment execution, the Portal provides a link to a Grafana dashboard for easy visualization of the data generated by the experiment.

Recommendation B5G-app VF Capabilities #12: Interfaces to Vertical Experimenters (portal data model)

For supporting all the features previously outlined, the Portal could store some information in its database. The main entities managed by the Portal are shown in Figure 5.

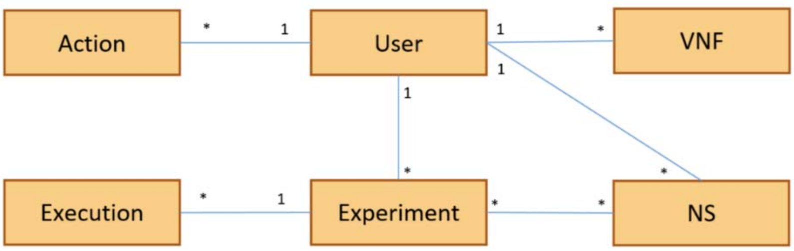

Figure 5: Entities of the Portal

The User Model (illustrated in Figure 6) contains the username, organization, email, and hashed password of a user. The password is saved encrypted so that not even Platform administrators can know this value by inspecting the database.

id	username	email	password_hash	organization
1	TestUser	test@5genesis.com	pbkdf2:sha256:50000$FlcGt[...]	5GENESIS

Figure 6: User model

The Experiment and VNF Models (shown in Figure 7) store the information required for running execution. The Experiment Model contains a reference to the experiment owner, the type of experiment (currently only standard experiments are supported), a list of test cases and UEs and a network slice identifier (optional). The list of network services used in the experiments is saved on an auxiliary table called 'experiment_ns' (not pictured) in order to support the many to many relations.

id	user_id	name	type	test_cases	ues	slice
1	1	TestExperiment	Standard	["RTT"]	["Galaxy S7"]	Test Slice

Figure 7: Experiment model

The VNF and NS Models (represented in Figure 8 and Figure 9 respectively) are very similar. Both include the name of the instance, a description and a reference to the owner. In the case of the VNF model, there are columns for saving the image and VNFD file names, while on the NS model the NSD file name is stored.

id	user_id	name	description	VNFD	image
1	1	Test VNF	This is a test VNF	test.vnfd	test.image

Figure 8: VNF model

id	user_id	name	description	NSD
1	1	Test NS	This is a test NS	test.nsd

Figure 9: NS model

The Execution Model (represented in Figure 10) stores the start and end time of an execution, the dashboard URL for the results page, the status and current percentage of the execution and a message that informs the user about the current execution step. The last three attributes are continuously updated during the execution and are displayed to the user in the dashboard.

id	experiment_id	start_time	end_time	dashboard_url	status	percent	message
1	1	2019-05-29 10:34:57.25	2019-05-29 10:36:17.28	/d/Run1/experiment-run-1	Finished	100	Finished

Figure 10: Execution model

The Action Model (shown in Figure 11) is used for recording the different actions that a user performs. This information is used for generating a feed that will be displayed in the user's dashboard. This feed provides direct access to the results of the latest experiment executions, for example. In order to generate this information, the Portal stores a reference to the user, the timestamp of the event and a HTML encoded message that includes the relevant links.

id	user_id	timestamp	message
1	1	2019-05-28 09:08:33.40	Uploaded VNF: TestVnf

Figure 11: Action model

Recommendation B5G-app VF Capabilities #13: Intent-based approach and AI/ML

Intent-based Networking (IBN) provides a new approach to networking, where planning, design, implementation and changes to the network are made automatically by using special software. IBN's main goal is to simplify the creation, management and policy enforcement to the network with the use of Artificial Intelligence (AI), network orchestration and Machine Learning (ML).

The application of an intent-based approach in the testing and validation of 5G and Beyond environment will have the following benefits for the system:

- Reduces the complexity of the management and maintenance of testing and validation procedures.
- Simplifies the deployment of network services.
- Reduces labour associated with traditional configuration of switches and routers.

- Strengthens network security capabilities.
- Improves agility of the entire network system especially for the execution of a set of tests with different configurations.
- Eliminates repetitive or error-prone coding associated with manual configurations.

The adoption of an intent-based approach in the testing and validation environment should incorporate the following four aspects:

- Translation and validation: The system can translate a given command or business intent into actions that the software can perform. Additionally, it verifies that the intent can be executed successfully in the first place.
- Automated implementation: Once the intent or desired state is defined, the system will allocate network resources and enforce policies to meet the goal.
- State awareness: The system will continuously gather and monitor data to reflect the current state of the network.
- Assurance and dynamic optimization/remediation: Using machine learning, the system will implement and maintain the desired state of the network, applying automated corrective action if necessary. ML gives the network the ability to analyse, extract and learn from data dynamically.

In a testing and validation environment with intent-based capabilities, the verticals can express their "intent" of test and validate their services in the 5G and Beyond network, without providing any configuration. Instead, the testing and validation framework should figure out what actions are necessary to be taken, to provide the expected network state (intent) to realize the tests. Having stated that, it is not requested from the vertical to have any basic knowledge about networking. In addition, the configurations are realized in an automatic way without to need for the administrators to configure every device needed to accomplish the tests.

The intent-based approach will provide a way for the vertical experimenters to instantiate an experiment. The intention of the user will be collected and translated into predefined templates (e.g. Blueprints) describing his/her intended service needs in networking terms, thus enabling the creation of experiment descriptions assisted via natural language.

4.3.2.2.3 Processes

Recommendation B5G-app VF Processes #01: Application Validation end-to-end workflow

The B5G-app validation framework should be designed in order to be used by different Vertical industries to run their experiments in a B5G enabled infrastructure to discover and evaluate the value-adding business potential of a B5G mobile network when applied to Vertical services. The execution of Experiments and associated test cases should be structured in four main phases as shown in Figure 12.

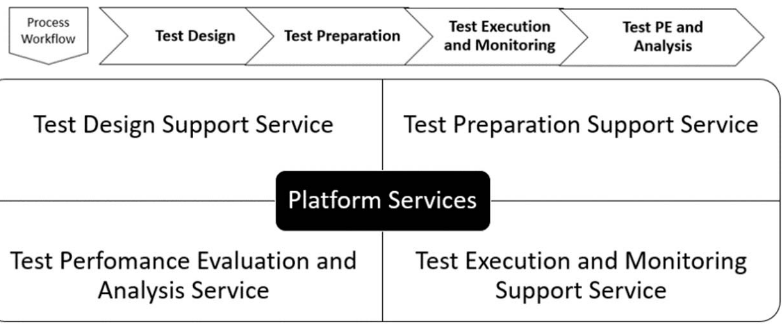

Figure 12: Application Validation end-to-end workflow

Along the experiment flow process, an active dialogue between Verticals and B5G-app platform stakeholders make possible to come to a full reciprocal understanding of B5G-app platform capabilities and Vertical use cases requirements and agree on the trial specifications including the use cases to be validated, scenarios covered, interfaces used, metrics and KPIs measured, etc.

This should then be materialized in experiments' design and development activities that will provide Verticals with a tailor-made 5G and Beyond virtual infrastructure to run their experiments. Experiment execution should support Verticals to validate their applications in customized scenarios to simulate the characteristics expected in real commercial deployments. This approach will allow Verticals to collect the measurements needed to fully assess the performance of their services in scenarios replicating realistic operational environments, with the final objective of tuning the configuration of their applications to meet the conditions and characteristics of different deployment options, thus maximizing the efficiency of the service roll-out phase.

In this sense, it is key to fully describe the Experiment Flow phases including:

- *What* - Activities included, inputs and expected outputs.
- *Who* - B5G-app VF platform main stakeholders and actors.
- *How* - Process and procedures.
- *When* - B5G-app VF platform availability.

The results achieved will help Verticals to build their business solutions in 5G and Beyond commercial networks implementing 5G and Beyond standards releases. The challenge is to provide a 5G and Beyond infrastructure that has the capacity, capability, reliability, availability, and security to support use cases defined by the Verticals using B5G-app VF as the experimentation platform to run their experiments.

Recommendation B5G-app VF Processes #02: Experiment requirement collection

The main goal of the Test Design phase is to understand what the Vertical needs concerning experimentation in a 5G and Beyond testing and validation environment. Although some tools exist related to test specification, normally these are more focused on making the life of the test design technician easier (e.g. time management applications, equipment catalogues, etc.) rather than on permitting a third party to express particular needs, like it is required in such case. Furthermore, the process is even more complex when considering that not all Verticals have the technical knowledge related to 5G networks, MEC, SDN, MANO, etc., as to be able to prepare a test plan on their own. Therefore, this is (still) a process very much based on human interaction, discussions, clarifications, etc., similar to the requirements specification phase on engineering projects, where adequate interaction with the client is key to the project success.

The information provided by the verticals should be used by the platform operators to configure the network scenarios that are relevant for the testing of the use cases defined by the vertical, the identification or definition of teste cases need to compute the KPIs, the customization of the experiment descriptor and the definition of network slice if applicable.

The scenarios, the experiment descriptor and the network slices will be the input to the testing framework, entity in charge of the execution of the experiment. The execution of the experiment involves the configuration of the experimentation platform, the control and execution of the experiment, the control and execution of the monitoring tool, the retrieval of the measurements and the generation of the final report with the results obtained.

In that sense, the generation of an experiment description template, to be used during the test design phase is of paramount importance. This template, called test plan template, should be circulated to the verticals willing to experiment using the 5G and Beyond infrastructure and should contains direct questions to the verticals in order to collect important inputs from them.

One important fact in the interaction between Verticals and Operators (or Manufacturers) is that the "language" is not always the same. Operators tend to focus on network parameters, conditions, SLAs, etc., while the implications of these concepts are not always fully understood by Verticals. Therefore, a lot of effort should be put in the template generation in order to make it understandable for experimenters, hiding the network related components as much as possible.

The template should be shared very early with the Verticals, to iteratively review the correctness of its content and to identify gaps or missing points. Once the template is finalized it will be shared with the Verticals.

A reference experiment description template may include the aforementioned information to be collected from the verticals:

- A brief description of the experiment, and to outline what capabilities the experimenter is seeking from the infrastructure.

- A list of the components that verticals will be bringing into as part of the experiment, together with their deployment requirements and how they interact with each other. This is important, on the one hand to help determining hosting capabilities for the experiment infrastructure, and on the other for understanding the required connectivity among components.
- List of the KPIs which are meaningful for the verticals and if these KPIs can be measured during experimentation.

This template will be the basic for the speciation of the test scenarios relevant for the verticals and the test cases used for the calculation of the KPIs.

Recommendation B5G-app VF Processes #03: Test Scenarios

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.).

A scenario includes information related to network and environment configurations and is related to the technologies supported by the experimentation platform/infrastructure. From the performance perspective, a scenario quantifies the parameters that affect the values of the KPIs to be measured.

Table 4: Scenario description template

Scenario Description Template	
#	**Description of the fields to be completed**
1	Radio access technology *4G, 5G*
2	Standalone / Non-Standalone (if applicable)
3	Cell Power
4	Frequency band: *Sub-6 GHz* *mmWave*
5	Maximum bandwidth per component carrier *50 MHz, 100 MHz, 200 MHz, 400 MHz*
6	Sub-carrier spacing *Sub 6 GHz: 15 kHz, 30 kHz, 60 kHz* *mmWave: 60 kHz, 120 kHz, 240 kHz, 480 kHz*
7	Number of component carriers *Maximum number of CC = 16 (5G)* *Maximum number of CC = 5 (4G)*
8	CP *Cyclic Prefix: normal, extended*
9	Massive MIMO *Number of antennas on NodeB*
10	MIMO schemes (codeword and number of layers) *The number of codewords per PDSCH assignment per UE* • *1 codeword for 1 to 4-layer transmission* • *2 codewords for 5 to 8-layer transmission.* DL DMRS based spatial multiplexing (SU-MIMO/MU-MIMO) is supported • *At least, the 8 orthogonal DL DMRS ports are supported for SU-MIMO* • *Maximum 12 orthogonal DL DMRS ports are supported for MU-MIMO*
11	Modulation schemes *Downlink: QPSK, 16 QAM, 64 QAM, 256 QAM* *UplinK: QPSK, 16 QAM, 64 QAM, 256 QAM*
12	Duplex mode *FDD, TDD*
13	TDD uplink/downlink pattern (if applicable) *0,5 ms, 0,625 ms, 1 ms, 1,25 ms, 2 ms, 2,5 ms, 5 ms, 10 ms*
14	Contention based random access procedure/contention free (if applicable)
15	User location and speed
16	Background traffic
17	Computational resources available

Recommendation B5G-app VF Processes #04: Test Cases

To ensure experiment repeatability, regardless of the test equipment and the entity performing the certification, experimentation processes have to be based on specific testing procedures, configurations and conditions. For that purpose, specific test cases have to be elaborated. Each test case apart from the description of the KPI under test and its measurable objectives should include:

i) the configuration of the testing environment (for all RATs, network components, software tools and hardware);

ii) the set of procedures to perform the measurement;

iii) the measurements points; and

iv) the formulas needed to calculate the KPIs.

A test case provides uniformity and organization to run an experiment in a programmatic and structured way. A potential example of Test case template is described in Table 5.

Table 5: Test case template

Test Case Template	
#	**Description of the fields to be completed**
1	Description of the target KPI *Here goes the definition of the target KPI. Each test case targets only one KPI (main KPI). However, secondary measurements from complementary KPIs can be added as well (see field 4 in this template). The definition of the main KPI specializes the related target metric (the ID of the related target metric is declared in the first row of this template). More precisely, the definition of the main KPI declares at least the reference points from which the measurement(s) will be performed, the underlay system, the reference protocol stack level, etc.*
2	Methodology *Here the acceptable values for the monitoring time, the iterations required, the monitoring frequency, etc., are declared. The reference to the calibration test is taken from the test case. This is to facilitate the comparison between measurements.*
3	*Calculation process and output* *Here goes information related to the calculation process required. This is information may include details related to the underlay system. Here goes also the Units of the metric, and potentially a request for first order statistics (Min, Max, etc.)*
4	Complementary measurements *A secondary list of KPIs useful to interpret the values of the target KPI. Getting these measurements is not mandatory for the test case.*
5	Pre-conditions Any requirement that needs to be done before *execution* of this test case. A list of test specific pre-conditions that need to be met by the SUT including information about **equipment configuration**, **traffic descriptor** i.e. precise description of the initial state of the SUT required to start executing the test sequence
6	Applicability A list of features and capabilities which are required to be supported by the SUT in order to execute this test (e.g. if this list contains an optional feature to be supported, then the test is optional)
7	Test Case Sequence *Specializes the measurement process (methodology) of the metric for the selected underlay system. Measurements points and measurement procedure specification.*

Recommendation B5G-app VF Processes #05: Experiment Descriptors

The formalization of the experiment if the step prior to its execution. The formalization should be based on the information provided by the experiment definition template, the test cases and the scenarios. The experiment descriptor should be agnostic to the testing equipment and to the entity performing the certification/validation. The testing framework will be in charge of the translation of the experiment descriptor to the final set of actions to be executed (test plan).

The formalization of the experiment is an important step for the assessment/comparison of the outputs obtained during the execution for the experiment. The experiment descriptor should be well-structured and formalized. To this end, all the required information for running the experiment is recommended to be included in the experiment descriptor.

A potential example of Experiment Descriptor template can be seen below:

```
{
  ExperimentType: Standard/Custom/MONROE
  Automated: <bool>
  TestCases: <List[str]>
  UEs: <List[str]>  UEs IDs

  Slice: <str>
  NSs: <List[Tuple[str, str]]> (NSD Id, Location)
  Scenario: <string>

  ExclusiveExecution: <bool>
  ReservationTime: <int> (Minutes)

  Application: <str>
  Parameters: <Dict[str,obj]>

  Remote: <str> Remote platform Id
  RemoteDescriptor: <Experiment Descriptor>

  Version: <str>
  Extra: <Dict[str,obj]>
}
```

The first two sets of values are the most important for the definition of the experiment. The first group includes the type of experiment, the test cases to execute and the UEs to use, while the second define the slice, network services and scenario to configure and deploy.

The third group is used to control the scheduling of the experiment. An '*Exclusive'* experiment will not be run at the same time as other experiments in the testbed, while the '*ReservationTime'* is used to define the duration of the experiment when automation is not enabled.

The fourth group is used to define the configuration existing experimentation solutions: '*Application'* defines the container to deploy in the node, and '*Parameters'* includes the configuration of the container. The '*Parameters'* field is also used for specifying customized parameters in the case of a 'Custom' experiment.

The fifth group is expected to provide the fields necessary to support the execution of distributed experiments. The 'Remote' field is used to identify the secondary platform that will be part of the distributed experiment, while '*RemoteDescriptor'* contains a JSON object in the same format as the main descriptor, but excluding the 'Remote' and '*RemoteDescriptor'* fields. This secondary descriptor contains the values required to configure the experiment execution in the remote platform.

In order to ease the addition of new functionality in the future two fields have been included: The '*Version'* field can be used to specify the exact version of the Experiment Descriptor, so that the lower layers can customize the handling of the descriptor according to any future modification while keeping compatibility with older descriptors. The '*Extra'* field can be used to add any kind of information. This can be useful, for example, for adding debug or tracing information, or as an easy way to support extra functionality without changing the format of the Experiment Descriptors.

The Experimenter configures the experiment descriptor, filling in the missing information of the variable parameters defined during the collection phase. The result of this customization produces a set of completed descriptors which embeds all the information needed to instantiate and run the experiment itself. In order to ease the process, The B5G-app VF should offer a wizard that guides the experimenter in the definition of the descriptors, indicating all the parameters to be configured according to the experiment design.

Recommendation B5G-app VF Processes #06: Testing framework

The software tool and the guidelines/rules for creating and executing test cases define a testing framework. The testing framework should provide:

1) a well-defined format to specify inputs/test conditions (experiment descriptor, scenario, VNFs, etc.) and expected outputs;

2) an interface to introduce the inputs;

3) the actual test execution environment; and

4) a mechanism to report results and the status of the execution.

The testing framework should ensure that the experiment environment is properly prepared and configured to enable the experiment execution in a given timeslot and for a given configuration.

Three mandatory components of the testing framework architecture have been identified. These components are:

- The Scheduler is responsible for managing the execution of the experiments on a high level: An experiment execution is divided in 3 stages (Pre-Run, Run and PostRun), and the Scheduler keeps track of the execution of each of these stages for multiple experiments in parallel.
- The Execution Engine includes the logic for managing the execution of each experiment stage, by generating an independent Executor. The progress in each Executor is further divided in different Tasks, which are dependent on the test case and the equipment involved in the experiment.
- The Composer is responsible for creating the Platform Specific Configuration of the received experiments. The configuration generated includes the Tasks to be run by the Executors.

The following components have been identified as optional components:

- The Administration Interface provides a unified interface to platform administrators where they can review the execution status of active experiment run, as well as checking the logs generated by every execution, including previous ones. From this interface it should be also possible to cancel the execution of an experiment.

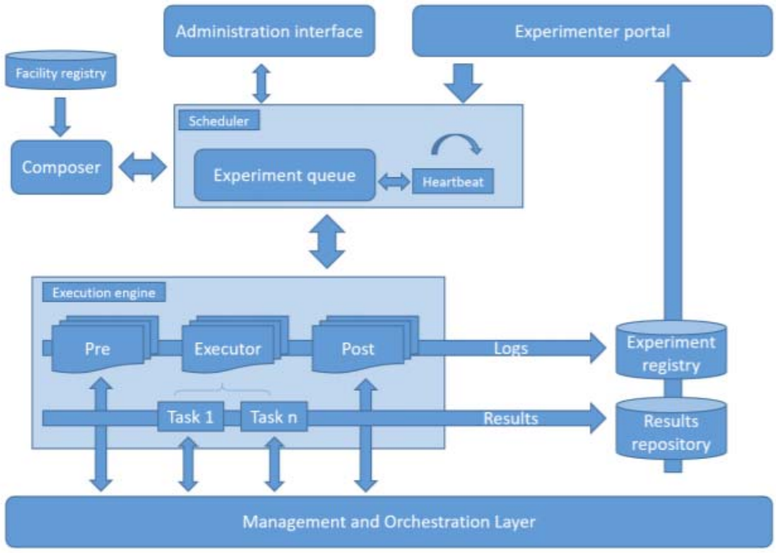

Figure 13: General architecture for the testing framework

The work-flow of the testing framework when an experiment execution is requested is as follows:

- The Scheduler creates a new Experiment Run instance. These objects contain all the information about a particular execution.
- The Scheduler requests the creation of a Platform Specific Configuration to the Composer, using the Experiment Descriptor received on the request.
- The Composer generates this configuration (including the Tasks to execute in each Executor).
- The Scheduler queues the experiment execution, starting from the Pre-Run stage. The execution is then handled by the Pre-Run Executor, which runs on a separate thread and will wait until all resources are available (among other actions).
- When the Pre-Run executor finishes (which means there are available resources), the Scheduler moves the experiment to the Run stage. Again, the real execution of the Tasks is handled by a different thread in parallel.
- The Scheduler moves the execution to the Post-Run stage once the Run stage finishes, and additional Tasks runs on the new Executor.
- When finished, the Scheduler removes the Experiment Run from the queue.

Recommendation B5G-app VF Processes #07: Test Execution and Monitoring process

The test execution and monitoring process involves the following steps:

- **Experiment instantiation:** firstly, all the resources and components related to the experiment to be executed should be deployed before starting with the execution of the test cases. This implies the request for the creation of the proper Network Slices associated to the given experiment, resulting in the instantiation of the NFV Network Service instances based on the requirements defined in that experiment. During this instantiation phase, the configuration of the components that perform functions related to monitoring and performance analysis is also provided and applied.
- **Experiment tests execution:** when all the resources related to the experiment are available in the testing platform, the corresponding applications and the scripts defined in the experiment specification will be executed according to the test plan. This phase can be also decomposed in several sub-phases:
  - A first common step may involve the configuration of the components involved in the experiment.
  - Once these components are properly configured, the testing system can start to execute the test cases and collect the related monitoring data (for this last point, check the following bullet point - Experiment monitoring).
  - As soon as the execution is terminated, apart from cleaning the configuration applied to the experiment components, the virtual environment associated to the experiment and the functions that carry out monitoring processes can be also cleaned, to release the related network and compute resources.
- **Experiment monitoring:** during the experiment tests execution, apart from monitoring the correct behaviour of the system under test, the Experiment metrics, KPIs and results defined in the test specification are collected through dedicated Experiment Monitoring tools, which extract that data from the components related to the experiment (VNFs, PNFs, etc.) and deliver it to a Monitoring platform that allows the verticals to check the achievement of the desired KPIs. This information can be also available after finishing the experiment.

Recommendation B5G-app VF Processes #08: Concise and self-descriptive reporting of validation results

The testing and validation environment for 5G and Beyond should include a complete and well-defined results analysis and validation reporting mechanism, which provides experimenters with clear answers to their initial validation questions. In addition, this reporting mechanism should give a complete picture of the testing and validation process realized using the infrastructure including, in addition to passed/failed results, information about the actual testing process, the environment in which the tests are executed, the different test cases executed to provide the results, a set of high-level configuration information. All this information can be provided in a way that will ensure the transparency and repeatability of the entire testing and benchmarking process.

In this direction, the final report should be consolidated in a composition of a set of sections created by the elements that participate in the testing and validation platform. This is important because, while the main focus of the experiments is the KPI validation, the information regarding the infrastructure, the conditions and the technologies used can be extremely useful and insightful to the experimenter as well.

Reporting the information generated from each of the various stages of the experiment definition, preparation and execution processes may happen independently in the respective stages, concluding to a complete self-descriptive report. The complete report may contain the following clauses:

- Test Case Validation clause: It pertains to the targeted KPI that the Vertical wants to validate containing information regarding the behaviour of the KPI throughout the test run as well as the final result of the validation process.
- Test Case report clause: It includes the results of the experiment operations returned by execution of the different test cases of the same vertical service and it is an operational report mainly focused on the different stages of the test execution process and less about the KPIs.
- Experiment clause: The present document contains all the information regarding the requested parameters, technologies used, use case details and other details pertaining to the experiment at a higher level.
- Scheduling clause: This information is produced by the GUI used by the experimenter, and as the name suggests, are all related to scheduling experiment executions like the time slot, the one or more selected sites and so on.

4.3.2.2.4 KPI Mechanisms

Testing and Validation activities, of any type, can only lead to meaningful and actionable conclusions when quantitative results can be actually collected, compared to expectations, and analysed. In the specific challenge of the test and validation of vertical applications over 5G and beyond networks this is even more evident and more relevant since end-to-end application behaviour, and as a consequence end-user experience, is enabled and conditioned by vertical application solution architecture and performance, the deployment model selected over the mobile network and the levels of performance granted the underlying network supporting it.

So, when specifying experimentation test cases (at unit and system levels) special attention should be dedicated to understand, identify and select performance indicators, at both application service and network service levels. And, depending on the specific testing objectives and strategies for a particular vertical application and its stage of development, the focus on vertical KPIs and network service KPIs can vary, from information to be monitored along the experiment execution process to targets set to be closely watched and validated. Given such context a B5Gapp-VF is expected to provide verticals with services fulfilling this broad range of needs by implementing robust and trustable mechanisms for KPI selection, collection, monitoring and analysis.

Additionally, the trend of gaining deeper insight into the characterization of the performance of vertical applications *in relation* with the observed and experience performance delivered by the supporting B5G network demands further attention. In order to assist the stakeholders of the validation process to be able to find out (or discard) and even quantify correlations between the vertical (application-level QoE) KPIs and the B5G network (5G capabilities level) service performance, a holistic approach incorporating tools and technologies for advanced data analytics and machine learning should be considered. This poses interesting expectations in to the KPI mechanisms and frameworks to be crafted in a B5Gapp-VF, which can be instrumental in securing technology-readiness under a number of varying target environments and conditions.

At this point, and before moving on to a general body of specific recommendations with a lower level of detail, it is worth considering a practical segmentation vertical stakeholders of the validation process into categories of common observed patterns of their expectations on experimentation, testing and validation, with regards to KPIs. Each of these categories will place expectations of different levels and ambitions on the B5Gapp-VF, helping to secure the best adequacy of the recommendations to a broad range of scenarios and needs. The basic framework of general stakeholders' expectations introduced in clause 4.1.2 is used for framing these categories of vertical experimenters along with their respective fundamental needs related to KPI management:

a) Application-network interoperability verification:

At this stage and ambition level of testing and validation, vertical stakeholders focus their experimentation on the verification of the key aspects of integration between the vertical solution implementation and the target B5G network capabilities supporting it. For assistance to such verification at least a cost-grained focus on the most relevant vertical KPIs and on the B5G network service KPIs with major influence in the type of service under test is advisable and sufficient. No special emphasis on quantitative analysis of KPIs of either nature is relevant at this stage as long as interoperability can be verified, but that said, the KPI collected data may very well used to confirm or discard interoperability issues and support some troubleshooting actions that are not rare to be needed at this stage. For that purpose, it is advisable that at least a few key performance indicators of the application are monitored at experimentation time and analysed for validating conformance.

Also, at network service performance level, regardless the specifics of the vertical application which interoperability with 5G is under evaluation, two main categories of services can be considered: throughput-sensitive and latency-sensitive services of communications). So, User Data Rate, and RTT Latency, respectively should become the major network service KPIs at focus at this stage.

Summarizing, for interoperability verification purposes, KPI collected data can shed light on aspects and issues to be overcome, and by considering them at this early stage, also valuable information for supporting the planning of formal validation activities at further stages is made available.

b) Application end-to-end performance evaluation:

At this scenario the level of ambition goes well beyond assuring conformance of interoperability of the application with the targeted B5G network setup, and into determining the expected performance of the application under a wide range of conditions. Along the experimentation process its key to collect objective data of the E2E application performance and compare them to expected levels known and set in advance (that is the ultimate goal of validation). And it is also very interesting, for comprehensive analysis purposes, to collect metrics on the B5G network service performance as well. At this stage, a fine granularity in both vertical KPIs and B5G network service KPIs is advisable. Vertical target KPIs are well-defined and can be monitored; network service KPIs to be monitored are well identified, and a good approximate idea of the ranges expected is also developed. A richer set of B5G network service KPIs (the 5G Core KPIs proposed by 5G PPP WG TMV) is considered at this stage, extending from User Data Rate and RTT Latency to Availability, Reliability, Mobility and even Connection Density. KPI data management shows more and more relevant, and actually guides the process on a data-driven fashion, from one validation iteration to another. The two following scenarios are considered, based on the outcome of the validation iteration:

- Case of vertical KPIs having been validated:

  The actual performance levels of the B5G network service when the vertical KPIs are validated become well known, and, thus, a *performance-safe* model for target KPIs on the B5G network can be inferred, and used for next validation iterations over the same B5G setup or others (see next item). This is a fundamental purpose of this stage since it is the effective transfer from experimentation results into inputs for the SLA negotiation on B5G to be done before commercial launch of vertical services.

- Case of vertical KPIs NOT having been validated:

  The actual performance levels of the B5G network service when the vertical KPIs are NOT validated become well known. That way, a root cause analysis for determining the actual feasibility of the application over the targeted B5G setup would extremely benefit from using and collating this data, and to establish -or discard- potential correlations with the measured vertical KPIs. This is fundamental for assessing whether or not the vertical application should be tuned or optimized, or another type of B5G network setup should be targeted.

c) Network technology assessment:

In this scenario a former and thorough characterization of application targeted and actual KPIs and the expected (target) network service KPIs, for the targeted B5G network setup has been carried out. The aim now is to explore alternative B5G setups, confirming or discarding the feasibility of the application over them, for which KPI analysis is the key tool, leveraging the KPI model inferred in the first completed validation.

The usual validation strategy for this purpose is to evaluate both vertical and network service KPIs vs the targeted values identified in former completed validation, so it can be a very convergent and agile process. It is also common practice that extra specific KPIs are considered since the comparison between alternative B5G network set-ups bases not only in previously identified performance dimensions but also on additional aspects that may have not been assessed yet (examples are energy or resource consumption in general, service deployment time, etc.) but that now can support making well-informed decisions.

d) Application deployment model optimization:

In this scenario also a former and thorough characterization of application targeted and actual KPIs and the expected (target) network service KPIs, for the targeted B5G network setup has been carried out. The goal now is to explore alternative deployment models over the same type of B5G setup, for confirming or discarding their feasibility. KPI analysis is a key tool for validation assessment and for benchmarking, leveraging the KPI model inferred in the first completed validation, and iterating over it. Same as in the former item, the usual validation strategy for this purpose comprises evaluation of both vertical and network service KPIs vs the targeted values identified in former completed validation. It is also common practice that a few extra specific KPIs are considered since the comparison between alternative application deployment models bases not only in previously identified performance dimensions but also on additional aspects that may have not been assessed yet (examples are service deployment time, fault-tolerance, etc.) but that now can support making well-informed decisions.

Despite the huge variability on performance validation needs and scenarios and the intrinsic heterogeneity of vertical use cases and solutions addressed in the present document, a common general model for dealing with KPI-assisted testing and validation of vertical application over 5G and beyond networks would be an asset of utmost relevance. And, therefore, a fair trade-off between versatility and general applicability is postulated here. A general, pragmatic, and purpose-fit approach is outlined in the series of recommendations on KPI management for a B5Gapp-VF listed here.

Recommendation B5G-app VF KPI Mechanisms #01: KPI Selection

It should be entirely up to the vertical stakeholder to select the specific set of KPIs to be monitored at vertical application and 5G network service level for serving its specific purposes and scenarios of validation. It should also be up to the vertical stakeholder to decide on the involved profile of action for each KPI, ranging from observing it to evaluating it vs target KPIs defined.

With this freedom of choice over a wide range of options, the vertical experimenter can collect valuable data for further analysis and gain insight on potential correlations on vertical KPIs and 5G network service KPIs, for a variety of testing conditions. This possibility is advised to be balanced with an exercise of analysis of major influencing KPIs, a priori, on the service performance, in order to converge to conclusions easily and fast, whenever possible. In particular it is recommended to be able to classify the addressed use case into one of these three categories (eMBB, URLLC and mMTC enabled services) so that concrete network service KPIs (respectively User Data Rate, RTT latency and Connection Density) can be especially focused from the beginning of the validation process. Expanding the analysis to secondary or further KPIs should be a natural step, whenever needed, and facilitated by the B5Gapp-VF KPI mechanisms and services.

In order to identify the subset of 5G Service KPIs of major influence to each identified vertical KPI, Table 6 can be used for guidance. It illustrates recurrent patterns of 5G Service KPIs for typical categories of B5G-enabled vertical applications.

For instance, for the category of streaming services, Table 6 allows to easily look up that:

a) KPIs Minimum Expected Throughput and Maximum Expected Jitter are the essential ones to monitor and analyse;

b) KPIs Minimum Expected Availability, Minimum Expected Reliability, and Maximum Expected RTT Latency are likely to be irrelevant; and finally

c) KPIs Minimum Area Traffic Capacity, Minimum Connection Density, and Minimum User Mobility could have or not an influence for some specific applications, which should be studied for your application.

Table Legend:

Y: Sensitive KPI

N: Irrelevant KPI

*: Sensitivity of this KPI is Use Case dependent

Table 6: 5G Service KPIs of B5G-enabled vertical applications

5G SLICE TYPE	VERTICAL APPLICATION CATEGORY	Minimum Expected Throughput (Mbps)	Maximum Expected Jitter (msec)	Minimum Expected Availability (%)	Minimum Expected Reliability (% for Lat.)	Maximum Expected RTT Latency (msec)	Minimum Area Traffic Capacity (Mbps/m^2)	Minimum Connection Density (sessions/m^2)	Minimum User Mobility (Km/h)
eMBB	Web Surfing	Y (DL)	N	N	N	N	*	*	*
	Download	Y (DL)	N	N	N	N	*	*	*
	Upload	Y (UL)	N	N	N	N	*	*	*
	Streaming	Y (DL)	Y	N	N	N	*	*	*
	UpStreaming	Y (UL)	Y	N	N	N	*	*	*
URLLC	Remote Config	N	N	Y	Y	Y	*	*	*
	Remote Control	N	N	Y	Y	Y	*	*	*
	Synched Devices	*	*	Y	Y	Y	*	*	*
mMTC	Telemetry	N	N	Y	N	N	N	Y	*
	Actuation	N	N	Y	N	N	N	Y	*
	Interaction	N	N	Y	Y	N	N	Y	*

Recommendation B5G-app VF KPI Mechanisms #02: B5G network service KPIs Support

For any type of 5G environment, capabilities, and configurations, and regardless the variability in that space, a common set of well-defined B5G network service KPIs should be considered and supported by any B5G-VF. That means that, for each B5G network service KPIs considered, the B5G validation framework is expected to implement the necessary mechanisms for collecting, monitoring and exposing it to the KPI Data Framework following the best practice established in EU 5G PPP projects established in [i.60].

Such set of B5G network service KPIs (in 5G PPP TMV [i.1] being called Core KPIs) includes:

- Service Availability.
- Mobility.
- Connection Density.
- Minimum Expected Upstream Throughput.
- Minimum Expected Downstream Throughput.
- Maximum Expected Latency.
- Minimum expected Network Reliability.
- Uplink Peak Throughput.
- Downlink Peak Throughput.
- Maximum Expected Network Jitter.

With this approach a guarantee of support, compatibility and conformance for those selected and recurring 5G network service KPIs is built-in in the B5G-VF, and without precluding the incorporation and support of other network KPIs on the same footage.

As referred before, 5G PPP TMV WG has recently released a white paper describing a mapping exercise between vertical/service KPIs and network/core KPIs in [i.42].

Recommendation B5G-app VF KPI Mechanisms #03: B5G network service KPIs Measurement Points

In addition, to support of a set of KPIs (at least the B5G network service KPIs) as described in the previous recommendation, it is also important for a B5G-app VF to define the appropriate observation points for each KPI, in which the mechanisms for KPI collection will be developed.

Observation points are identified as key locations to correctly measure the values from the B5G network service KPIs. Following recommendations are based on the observation points defined in 5G PPP TMV [i.1] these observation points can be further classified in two mayor groups, the measurement points located at the 5G system (further referenced as 5GS) and the points located close or at the vertical applications (APP E2E):

- **5GS observation points** are limited to interfaces within the 5G domain, including the 5G Radio, Edge, Transport and 5G Packet Core, the KPI data availability may be dependent on the available data collection methods such can range from external probing systems, OAM data or network exposure data. KPIs measured in this domain can be used to derive the corresponding end-to-end KPIs or even be used to validate the correctness of end-to-end measurements at the endpoints, they are exclusively related to the domain, and are not designed to provide end-to-end information.
- **APP E2E observation points** provide end-to-end information about the service behavior, but they are limited to the applications and services the vertical controls or owns. This observation point provides greater flexibility than the 5GS observation point, as verticals are not limited by the technical and privacy issues affecting the vendors and operators. In some cases, verticals are unable to deploy probes at the APP E2E observation points, for this case measuring at the 5GS could be a fallback solution, taking into consideration its limitations.

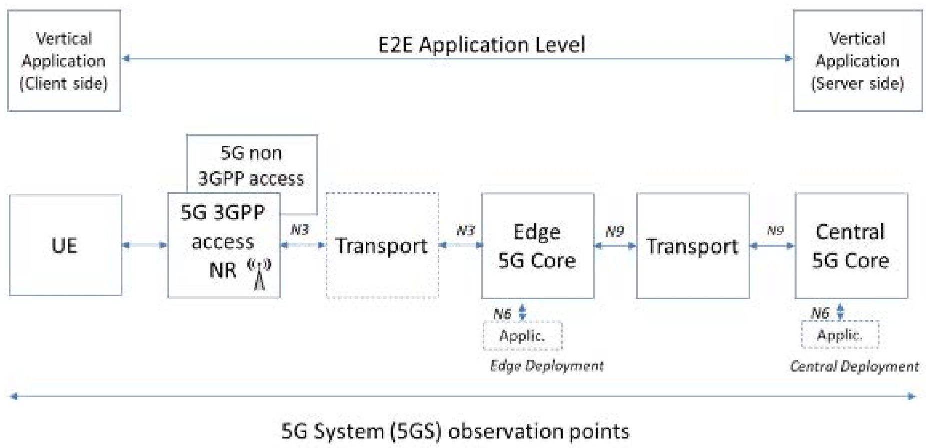

Figure 14: 5GS observation points and APP E2E observation points

As each B5G network service KPIs can have one or many observation points, in the following it is detailed a number potential observation points and cases of applicability for each B5G network service KPI.

B5G network service KPI 1 - Minimum Expected Upstream throughput

Table 7: B5G network service KPI 1 - Potential observation points

Observation Point Reference	Category	Location
KPI 1-OP 1	5GS	5G Core N6 interface endpoint
KPI 1-OP 2	APP E2E	N6 interface endpoint

The minimum expected upstream throughput KPI can be observed from two points, KPI 1-OP 1 and KPI 1-OP 2, both observation points take measurements from the same interface N6 but from different categories. However, taking in consideration what has been discussed earlier, measuring at the 5GS, is recommended solely if there is no possibility of measuring at the APP E2E.

B5G network service KPI 2 - Minimum expected Downstream throughput

Table 8: B5G network service KPI 2 - Potential observation points

Observation Point Reference	Category	Location
KPI 2-OP 1	APP E2E	UE interface with the 5G Radio endpoint

The minimum expected downstream throughput KPI can be observed from only one point, KPI 2-OP 1. It is recommended to only measure at KPI 2-OP 1, as measurements at other locations will not be as precise. The downlink throughput is affected by every component in the service chain, so for efficiency reasons it is only recommended to observe at the final endpoint (UE).

B5G network service KPI 3 - Maximum expected latency

Table 9: KPI 3 - Potential observation points

Observation Point Reference	Category	Description
KPI 3-OP 1	5GS	5G Core N6 interface endpoint
KPI 3-OP 2	APP E2E	N6 interface endpoint
KPI 3-OP 3	APP E2E	UE interface with the 5G Radio endpoint

The maximum expected latency KPI may be defined as round Trip Time and/or One way delay depending on the vertical application need and can be observed from three points, KPI 3-OP 1, KPI 3-OP 2 and KPI 3-OP 3. It is recommended to measure this KPI at the application (KPI 3-OP 2) or at the UE endpoints (KPI 3-OP 3). Measuring at the 5GS, is recommended solely if there is no possibility of measuring at any of the two other APP E2E observation points. In case the vertical is able to choose between KPI 3-OP 2 and KPI 3-OP 3, it is recommended to measure at the UE, as most traffic patterns originate at the UE.

B5G network service KPI 4 - Minimum expected Network reliability

Table 10: B5G network service KPI 4 - Potential observation points

Observation Point Reference	Category	Description
KPI 4-OP 1	APP E2E	N6 interface endpoint
KPI 4-OP 2	APP E2E	UE interface with the 5G Radio endpoint
KPI 4-OP 3	5GS	Edge N3 interface
KPI 4-OP 4	5GS	5G Core N3 Interface
KPI 4-OP 5	5GS	5G Core N6 Interface

The network reliability KPI can be measured from five points, KPI 4-OP 1, KPI 4-OP 2, KPI 4-OP 3, KPI 4-OP 4 and KPI 4-OP 5. It is recommended to measure this KPI at the application (KPI 4-OP 1) or at the UE endpoints (KPI 4-OP 2). Measuring at the 5GS, is recommended only if there is no possibility of measuring at any of the two other APP E2E observation points.

B5G network service KPI 6 - Uplink Peak throughput

Table 11: KPI6 - Potential observation points

Observation Point Reference	Category	Location
KPI 6-OP 1	5GS	5G Core N6 interface endpoint
KPI 6-OP 2	APP E2E	N6 interface endpoint

Similarly, as in KPI 1, the Uplink peak throughput can be observed from two points, KPI 6-OP 1 and KPI 6-OP 2. As stated in KPI 1 recommendation, measuring at the APP E2E is the recommended option. Measuring at the 5GS is recommended solely if there is no possibility of measuring at the APP E2E. So, it is recommended to observe at KPI 6-OP 2 instead of KPI 6-OP 1.

B5G network service KPI 7 - Downlink peak throughput

Table 12: KPI 7 - Potential observation points

Observation Point Reference	Category	Location
KPI 7-OP 1	APP E2E	UE interface with the 5G Radio endpoint

Similarly, as in KPI 2, the downlink peak throughput can be observed from one single point, KPI 7-OP 1. It is recommended to measure only at this point to maximize the precision and reduce the error on the measurements.

Recommendation B5G-app VF KPI Mechanisms #04: KPI Data Framework

A KPI Data Framework catering for generic mechanisms of KPI collection, storage, analysis and visualization should be a key component of any B5G-VF. With a domain-agnostic approach to KPI data management it is possible to leverage standard practices and tools, and secure flexibility, scalability and extensibility of the KPI analysis framework.

The KPI Data Framework mechanism intends to cover the full life cycle of the different kinds of KPIs and their related metrics during the execution of a given experiment. Such data can be related to network infrastructure or to applications, so that the service's performance can be evaluated at different levels.

The main responsibilities to be fulfilled by this process are:

i) collect data from different experiments and sources, logging it in a homogeneous way and doing a preliminary data manipulation;

ii) save the pre-processed data and perform different filtering operations in order to be easily analysed afterwards; and

iii) display the selected data for validation by either a human operator or an interoperating information system.

This results in the composition of a specific toolchain to deal with this KPI mechanism, in which each of the responsibilities identified before can be associated to a given entity:

- First of all, a **data collection, aggregation and pre-processing entity** would extract the experiment results, metrics and KPIs generated for a given experiment. The main challenge to achieve in this process is the management of multiple and different sources of monitoring data, such as activity log files, configuration data, active/passive probes or monitoring devices, among others. Each source would provide the data in a specific format, needing the application of data processing techniques to homogenize the data format to be received by the Monitoring system. Moreover, these sources may be located in different physical locations, so a distributed system is required to ingest the monitoring data from different sources. Some of the functions that can fit in this entity are:
    - Ingest data securely from multiple input sources simultaneously.
    - Support clock synchronization. The use of synchronized clocks between the different components of the environment is a good thing to have when retrieving data (such as metrics) for a given experiment. It is particularly important in the case of specific network metrics, for instance one-way latency, that use probes running at two different locations and need the synchronization of the clocks in order to give precise results. The B5G-app validation framework should make sure that the clocks and synchronization are precise enough so that they do not impact the validity of the measurements.
    - Execute different transformations and enhancements to the collected data by using filters, which parse each event, identify named fields to build structure and transform them to converge on a common format for more powerful analysis and business value.
    - Ship the data to various supported output destinations.
    - Extend and improve the previous pipeline with new plugins, which can be connected through specific APIs.
    - Guarantee at-least-once delivery for the data received with a persistent queue in case of failure, and also provide scalability to ingestion spikes without having to use an external queueing layer.

- Then, a **data indexing and storage entity** could search and filter among all the data gathered by the data collection entity for obtaining the useful information to be displayed, according to the results expected to be visualized during and after the experiment execution. A set of functions that can be offered by this entity are:
    - Provide a search and analytics engine with data storage capabilities.
    - Allow to perform and combine many types of searches (structured, unstructured, metric, etc.), and also performing data aggregation in order to explore trends and patterns in the data.
    - Leverage and access to all managed data at a very high speed thanks to the use of indexes for saving data.
    - Scale horizontally if needed, going from prototype to production seamlessly.
    - Rank the search results based on a variety of factors (from term frequency or recency to popularity and beyond). Mix and match these along with functions to fine tune how the results show up to the experimenters.
    - Detect failures to keep the deployed environment and the data safe and available with cross-cluster replication, using a secondary cluster as a hot backup.
    - Allow to connect, build and maintain clients in many languages such as Java, Python, .NET, SQL and PHP through the usage of standard RESTful APIs and JSON.
- Finally, a **data visualization** entity would be in charge of presenting those experiment results, metrics and KPIs with an intuitive GUI, being able to monitor the progress of the experiment in terms of that information displayed and allowing verticals to interact partially with the visualization tool in an online fashion (e.g. by choosing what kind of information they want to see in any moment, set-ting thresholds for some parameters, etc.). Some functions to be carried out by this entity are the following:
    - Visualize the data ingested in the platform.
    - Provide freedom to select the way to give shape to the data, using a huge variety of interactive visualizations.
    - Share visualizations to other actors easily by using the sharing option that works for each stakeholder (e.g. embed a dashboard, share a link, or export to PDF, PNG or CSV files and send as an attachment).
    - Organize the dashboards and visualizations through specific spaces.
    - Use role-based access control to invite users to certain spaces (and not others), giving them access to specific content and features.
    - Monitor the whole toolchain, enabling the configuration of additional features by using a visual UI.
    - Customize the way of representing data with unique logos, colours and design elements, uploading these designs to the platform in order to use them.

It is encouraged that the functions used for these different features are run as virtual functions, and deployed on demand when they are needed.

Recommendation B5G-app VF KPI Mechanisms #05: Vertical KPI model

For each vertical application to be validated an explicit model and set of Vertical KPIs (i.e. meaningful in the vertical domain itself) may and should be specified, and used as the fundamental basis for its actual formal validation. For each and every considered Vertical KPI, the vertical stakeholder is expected to implement the mechanisms for collecting, monitoring and, if deemed necessary, exposing them to the KPI Data Framework. With this approach a clear separation of concerns between the nature and collection of the heterogeneous vertical KPIs (responsibility of the vertical) and a generic KPI data collection and logging common for the various KPIs considered at any levels (responsibility of the B5G-VF) is achieved. That allows to integrate vertical service performance information into the analysis of the end-to-end system performance.

VERTICAL APPLICATION KPI	5G SERVICE KPI-1	5G SERVICE KPI-2	5G SERVICE KPI-3	...
V.KPI-A = <Value>			-	
V.KPI-B = <Value>		-		
V.KPI-C = <Value>	-	-	-	
...				

Figure 15: Vertical to 5G Service map

The recommended staged process for being able to tracking and formally modelling the real influence model of B5G network service KPIs on the Vertical-level KPIs -map structure illustrated in Figure 15, is as follows:

1) Identify the meaningful Vertical KPIs (V.KPI-A, V.KPI-B, etc.) at the vertical application domain level itself, defining the known limits for well-functioning and proper behaviour of the vertical application. This is responsibility of the vertical stakeholder, who has the expertise on its domain and application. It is extremely important to come down to a quantitative specification of the target levels (or, at least, ranges) for each of these Vertical KPIs, otherwise it will not be possible to extract meaningful objective conclusions from the experimentation process.

2) Secure the implementation of the metric systems in charge of monitoring and collecting the values of these vertical KPIs, and the integration with/in the B5G-VF for the involved testing and validation campaigns on this vertical application. This is a vertical stakeholder driven process, to be supported by the necessary B5G-VF services for the integration in the same KPI framework, so that the collected KPIs can be logged for further analysis *together wit*h the collected 5G KPIs.

3) Vertical KPI by Vertical KPI, the set of specific 5G service KPIs deemed of major, a priori, influence on each KPI. Those selected KPIs should, thus, be closely monitored in the testing and validation campaigns. The B5G-VF should support that selection and see to that at testing time such KPIs are monitored, collected and logged.

4) Whenever possible, a theoretical model for the target values (or, at least, ranges) of such influencing 5G KPIs should be produced. The B5G-VF should provide means for specifying those targeted values (or ranges) and secure their real-time monitoring for the hosted test and validation campaigns of that vertical application. This initial model can be a great input to the whole process, and may very well be refined over time based on actual data collected and gained insight along the process. It is important not to try to stick to the initial model, but instead open up for *finding out* the right model.

5) Once the tests are executed a comprehensive report, including information on both vertical and 5G KPIs, is created by the B5G-VF, and becomes accessible to both vertical and CSP stakeholders. The quantitative information included in such report can be instrumental for defining a further cycle of this staged process, leading hopefully to a converged model of mapping of vertical and 5G KPIs, an indispensable piece of information for all stakeholders to motivate technical and business decisions moving forward.

Recommendation B5G-app VF KPI Mechanisms #06: KPI validation

For the KPI evaluation process, three major pillars are considered namely, the statistical analysis, the measurements correlation check (in the case that in addition to the target KPI complementary measurements are available) and the prediction analysis based on ML tools.

- Statistical analysis

  The analysis of the KPIs is based on the results collected in the conducted experiments. An experiment consists of one or more test cases, which contain multiple iterations of a single test. The statistical properties of a single test are calculated from the measurements collected in the test. The statistical properties of a complete test case are obtained by taking the average of the corresponding properties of the test iterations in the test case. This results in a collection of normally distributed test case results whose averages will be close to the real value of the statistical property. Furthermore, it allows to specify confidence intervals for them using the Student-T distribution.

- Correlation and causality analysis

  Correlation is a statistical association between observed variables. Correlation techniques can reveal similar, or in extreme cases identical, behaviour between KPIs or other monitored variables. A high positive correlation is the most intuitive case, where two variables exhibit the same nature of change (increase and decrease). However, if two variables show opposing change (one always increases while the other decreases), they are also similar. This is called negative correlation. In a practical example, correlation analysis allows to test how strongly two KPIs are depending on a third KPI. A low correlation score might indicate that factors other than the monitored variables should be considered.

- Prediction analysis

  The focus here is on predicting how a change in the system might influence a variable's behavior. For example, a use case for prediction can be to estimate the required network deployment (e.g. adding or removing network elements) to guarantee acceptable KPI values. Other potential use cases include prediction of KPI degradation upon network element (re-)configuration and prediction of the effect of specific KPIs.

4.3.2.2.5 Vertical Applications Design

This clause identifies some recommendations related to the design of the vertical applications composing the services to be experimentally validated and evaluated in a B5G trial environment. They can be intended as general guidelines for software developers to design and implement 5G/B5G-enabled vertical applications compatible with the B5G-app Validation Framework mechanisms and able to fully exploit all its functionalities and benefits. Such recommendations do not need to be considered as mandatory, but as nice-to-have features that simplify their evaluation procedures and improves the efficiency and automation of their deployment and testing workflow.

Recommendation B5G-app VF Vertical Applications Design #01: Virtualized deployment

Vertical applications should be natively designed to run in virtualized environments, e.g. based on containers or Virtual Machines, exploiting the capabilities of edge and cloud environments in terms of flexible instantiations, variable dimensioning, sharing of virtual functions and resources, on-demand and/or automated scaling and migration. The virtualization of the services would make them fully compatible with the lifecycle management mechanisms provided by the B5G-app Validation Framework to automate their instantiation on the target 5G virtual infrastructure, at the edge or cloud sites. Moreover, the virtualized format would allow to experimentally validate the application behaviour and performance with variable flavours of deployment, for tuning the service configuration on the basis of different contexts and deployment targets.

Recommendation B5G-app VF Vertical Applications Design #02: Support for dynamic and programmable configuration

Vertical applications should be designed to enable their dynamic configuration upon the trigger from an external entity, using a secure connection. This feature would allow a smooth integration with the configuration mechanisms available as part of the lifecycle management procedures for the provisioning and day-2 configuration of services in the B5G-app Validation Framework virtual infrastructure. The support of programmable APIs to enable the VNF configuration would be an additional benefit to further simplify this kind of interaction.

Recommendation B5G-app VF Vertical Applications Design #03: Dedicated management interface

Vertical applications should be designed to expose a dedicated management interface towards the B5G-app Validation Framework. This interface would enable the communication between the B5G-app Validation Framework and the virtual components of the vertical application for VNF configuration, collection of monitoring data on vertical KPIs as generated from the application itself and trigger of actions for test execution purposes. The management interface may be exposed by one or more components within the vertical application, as required by the internal structure and configuration or monitoring procedures of the application itself. Moreover, such interface should be exposed through one or more external connection points of the vertical application, in order to make it reachable, in a secure manner, from the system components of the B5G-app Validation Framework.

Recommendation B5G-app VF Vertical Applications Design #04: Support for vertical KPIs measurement and collection

If vertical KPIs are relevant for the evaluation of the service performance, the vertical applications should implement internal procedures for the measurement and collection of such KPIs and integrate mechanisms to make them available on the KPI framework of the B5G-app Validation Framework. This would enable the processing of vertical KPIs in the B5G-VF procedures of the testing and validation campaigns, for the analytics and diagnostics elaboration. Optionally, the vertical applications may implement translation procedures to guarantee the compatibility between the internal format of the monitoring data and the common information model adopted for monitoring and KPIs at the B5G-app Validation Framework, to further simplify the integration with its monitoring system. Finally, the possibility to configure dynamically the type of vertical KPIs to be monitored and collected at the virtual application, as well as their frequency and aggregation level, may further improve the flexibility of testing workflow, giving the experimenter the opportunity to select on-demand the KPIs of interest.

5 Conclusion

The deployment of vertical applications and the need of assuring their expected performance and behaviour makes clear the necessity for vertical industries to experiment and pilot their "5G enabled" business cases before moving to the commercial stage. A clear advantage for all the stakeholders in 5G business is the definition of a common, generic 5G and beyond application testing and validation framework which validates the vertical application or service in a systematic manner under different 5G technology choices and deployment environments. The objective is also to get the vertical involved in the design and result evaluation phase, which goes beyond current network testing paradigm of service providers.

The present document surveys and reviews existing methodologies for testing and validation of vertical applications designed for leveraging the potential of 5G & Beyond networks, in order to identify existing gaps in such methodologies and propose solutions to cover them. The present document describes a Reference Testing & Validation Process, as well as Architecture, Point of Control and Observations, KPI validation strategies and mechanisms, and other aspects involved in the testing and validation of innovative vertical applications enabled by 5G & Beyond networks. The scope of analysis takes into consideration the vision, requirements, architectures and novel use cases of the 5G ecosystem.

Several recommendations are provided, as summarized in Table 13. They consider a number of aspects in terms of capabilities to be supported (e.g. 5G NR, edge and cloud capabilities, etc.), architectural approach (multi-site environment, capability exposure, etc), tools (for KPI collection and analysis, performance diagnostics, reporting, etc.), processes (experiment requirement collection, descriptors, etc.), and finally, a generic testing framework.

Table 13: Summary of recommendations

Aspect	Sub-aspect	Recommendation	Recommendation ID
Deployment and reference model		Principles of the reference model for the experimental platform	B5G-app VF Deployment #01
Capabilities	5G Capabilities & Enablers	5G NR capabilities	B5G-app VF Capabilities #01
		Network Slicing capabilities	B5G-app VF Capabilities #02
		Cloud and Edge Computing capabilities	B5G-app VF Capabilities #03
		NFV&SDN Capabilities	B5G-app VF Capabilities #04
		Interconnection and Extension capabilities	B5G-app VF Capabilities #05
		B5G Exposure Capabilities	B5G-app VF Capabilities #06
	Testing and Validation Environment	Application Deployment Environment	B5G-app VF Capabilities #07
		Application Testing Environment and Tools	B5G-app VF Capabilities #08
		5G Network KPIs and Metrics Framework	B5G-app VF Capabilities #9
		Performance Diagnostic capabilities	B5G-app VF Capabilities #10
		Interfaces to Vertical Experimenters - portal architecture	B5G-app VF Capabilities #11
		Interfaces to Vertical Experimenters - portal data model	B5G-app VF Capabilities #12
		Intent-based approach and AI/ML	B5G-app VF Capabilities #13
Processes		Application Validation end-to-end workflow	VF Processes #01
		Experiment requirement collection	VF Processes #02
		Test Scenarios	VF Processes #03
		Test Cases	VF Processes #04
		Experiment Descriptors	B5G-app VF Processes #05
		Testing framework	B5G-app VF Processes #06
		Test Execution and Monitoring process	B5G-app VF Processes #07
		Concise and self-descriptive reporting of validation results	B5G-app VF Processes #08
KPI Mechanisms		KPI Selection	B5G-app VF KPI Mechanisms #01
		B5G network service KPIs Support	B5G-app VF KPI Mechanisms #02
		B5G network service KPIs Measurement Points	B5G-app VF KPI Mechanisms #03
		KPI Data Framework	B5G-app VF KPI Mechanisms #04
		Vertical KPI model	B5G-app VF KPI Mechanisms #05
		KPI validation	B5G-app VF KPI Mechanisms #05
Vertical Applications Design		Virtualized deployment	B5G-app VF Vertical Applications Design #01
		Support for dynamic and programmable configuration	B5G-app VF Vertical Applications Design #02
		Dedicated management interface	B5G-app VF Vertical Applications Design #03
		Support for vertical KPIs measurement and collection	B5G-app VF Vertical Applications Design #04

These recommendations are expected to serve as guidance for a further specification of a B5G validation framework for vertical applications.

Annex A: EC 5G PPP platform solutions

A.1 5G PPP 5G EVE Platform

The 5G EVE platform offers an integrated set of tools, automated procedures and site facilities to allow Vertical industries to run their experiments in a 5G enabled infrastructure, providing virtual testing environments easy to customize in terms of contexts, test cases, metrics and KPIs to be collected, etc., where vertical services can be validated in realistic scenarios. The 5G EVE Testing as a Service approach is built around four pillars:

- automation of 5G experiment executions in customizable 5G virtual infrastructures dynamically deployed in intra- or inter-facility environments;
- monitoring of infrastructure and service metrics;
- collection and validation of KPIs;
- performance diagnosis.

The trial specification is the result of a close collaboration between Verticals and 5G EVE Platform stakeholders. Exploiting the 5G EVE platform capabilities and analysing the requirements of the Vertical use cases, the experiments' co-design and co-development activities define tailor-made 5G virtual infrastructures to run the Verticals' experiments in scenarios and environments with similar characteristics to those of real 5G standards-based commercial deployments to come. This methodology allows to collect application-level and infrastructure-level KPI measurements and fully assess the service performance in scenarios replicating realistic operational environments, as a step to properly tune the service configuration according to conditions and characteristics of various target deployments.

The execution of Experiments and associated test cases are structured in four main phases, as shown in Figure A.1.

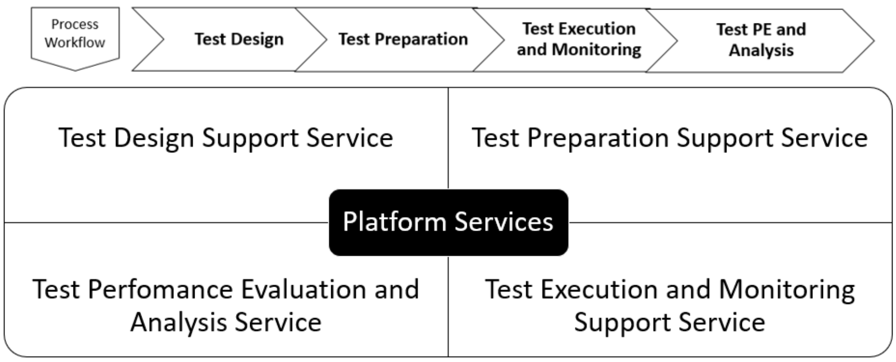

Figure A.1: Phases of 5G EVE experiment validation process

Experiment Design: In this phase, Vertical and other specialized actors, like VNF provider and Experiment developer, cooperate to identify the major characteristics, objectives and KPIs of the experiment related to a vertical service. The experiment specification is modelled through an experiment blueprint, which includes details about the vertical service under test and its deployment in the virtual environment, the operational context to run the experiment, the test cases to be executed, the metrics and KPIs to be collected and evaluated.

Experiment Preparation: In this phase, the experiment environment is properly prepared and configured to enable the experiment execution in one or more 5G EVE site facilities, during a given timeslot and for a given configuration.

Experiment Execution and Monitoring: In this phase, the dedicated virtual environment to run the experiment (and associated test cases) is built and configured, and finally the experiment Test Cases are executed. Metrics and KPIs are automatically collected, stored in the 5G EVE monitoring platform and elaborated.

Experiment Results Evaluation: In this phase, the experimenter analyses the experiment / test cases results.

The 5G EVE platform provides a multi-site facility to run the trials, with 5G-enabled testing infrastructures available in Italy, France, Spain and Greece. All the sites offer a variety of capabilities to test the service with different deployment environments. Details about the deployment of NSA and SA 5G architectures, spectrum and access technologies, support for network slicing, edge computing and NFV capabilities in each 5G EVE site are reported in the white paper "5G network support of vertical industries in the 5G Public-Private Partnership ecosystem" [i.72]. Moreover, the 5G EVE facility can be extended with additional sites, exploiting the features of its Inter-Working Framework.

The 5G EVE Portal offers a single point of access to all the 5G EVE functionalities for the design, execution and assessment of experiments. Using the web portal experimenters are able to define their experiment blueprints, configure and launch their experiments, verify metrics and KPIs during the experiment execution and visualize result reports with KPI evaluation and performance diagnostics information.

A.2 KPI collection and validation in 5G EVE Platform

The 5G EVE Platform offers mechanisms to automate the collection of metrics and KPIs for the validation of the vertical experiments. In particular, the experiment blueprint can declare two different types of metrics, i.e. application and infrastructure metrics, which are collected from the service applications or from the 5G network infrastructure deployed in each site, respectively. Infrastructure metrics are measurements of network performance or infrastructure usage and include RTT latency, user data rate in uplink or downlink, etc. Application metrics are service-specific measurements, like number of active sessions, number of connected users, etc. KPIs can be computed as functions of these elementary metrics and they can be evaluated automatically on the basis of thresholds defined in the experiment blueprint (see Figure A.2).

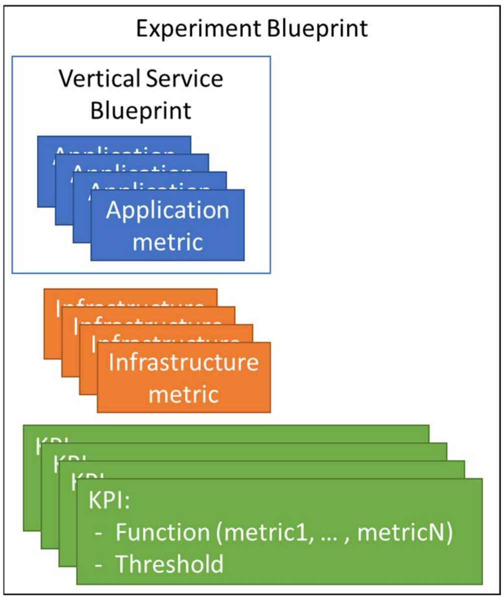

Figure A.2: Declaration of application metrics, infrastructure metrics and KPIs in 5G EVE experiment blueprint

Each 5G EVE facility deploys a number of network probes that are used to automatically collect the infrastructure metrics specified in the experiment blueprint. The collection of the application metrics, on the other hand, should be enabled by the service functions provided by the vertical.

In particular, application metrics should be generated by the vertical service itself using its internal, proprietary mechanisms (e.g. logs processing, application probes, etc.) and they need to be published into the Kafka broker of the 5G EVE monitoring platform using service-specific data shippers (e.g. https://www.elastic.co/beats/). Therefore, the application developer needs to implement the data shippers dedicated to the application metrics of interest and install them in the VM images that are used to instantiate the VNFs of the service. At runtime, these data shippers are automatically configured and launched as part of the experiment execution procedure and the application metrics are collected into the 5G EVE monitoring platform, where they can be processed and validated together with the infrastructure metrics.

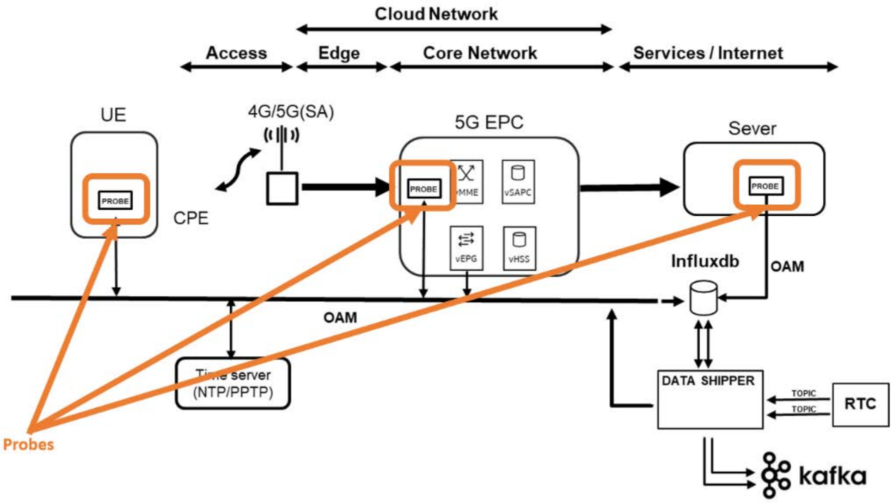

Figure A.3: Example of probes placement for collection of infrastructure and applications metrics in 5G EVE

An example of placement for probes dedicated to the collection of infrastructure and application metrics is depicted in Figure A.3. As shown in the picture, probes can be placed in the User Equipment, in the network infrastructure or embedded in VNFs running on the servers of the 5G EVE facilities. Their time synchronization is fundamental to guarantee the correctness of the measurements.

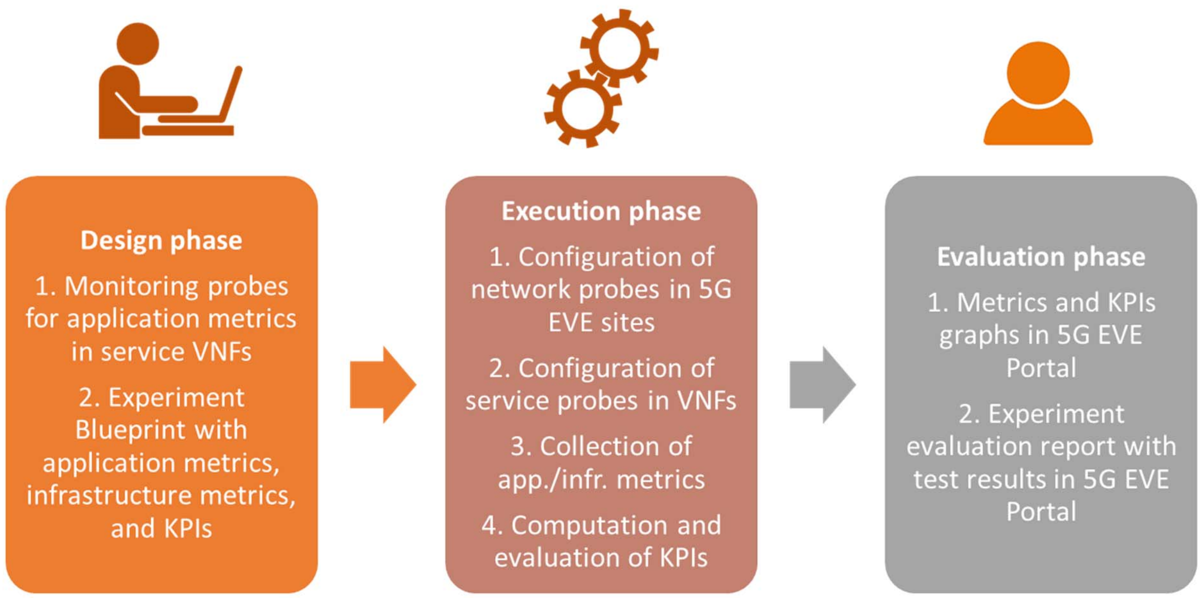

Figure A.4: Methodology for metrics collection and evaluation in 5G EVE experiment lifecycle

The overall workflow for the definition, collection and evaluation of metrics and KPIs in the 5G EVE framework from the perspective of the users is represented in Figure A.4 and can be summarized as follows:

Experiment design phase: The experiment developer identifies the infrastructure and application metrics required to evaluate the performance of the vertical service, how they can be combined together into relevant KPIs and the criteria to assess the service performance, i.e. the thresholds to validate the KPIs' values. For each application metric, suitable data shippers should be developed and installed in the VNFs' VM images. During the creation of the experiment blueprint, the experiment developer uses the wizard available in the 5G EVE portal to specify the application metrics (as part of the vertical service blueprint), selects the infrastructure metrics to be collected in each site and defines the criteria to compute and evaluate the KPIs. The configuration scripts of the test case blueprint can be used to further configure the data shippers for the application metrics collection, if needed.

Experiment execution phase: Starting from the information provided in the experiment blueprint, the 5G EVE platform executes the test cases selected by the experimenters and coordinates all the procedures for the collection, processing and evaluation of metrics and KPIs. In detail, the system creates the topics dedicated to the experiment metrics and KPIs in the monitoring platform and configures the network probes in the 5G EVE sites and the service probes in the VNFs. The metrics are collected through the Kafka broker and stored in the Data Collection Storage. The KPIs are computed in real-time, pushed in the monitoring platform and evaluated on the basis of the thresholds defined in the blueprint, producing a record with the overall results at the end of the experiment execution.

Experiment evaluation phase: The experimenter can analyse the results of the experiment through the 5G EVE portal. In particular, the graphs of metrics and KPIs can be visualized in dedicated dashboards, where they can also be filtered as needed. The values of metrics and KPIs can be downloaded in CSV files, e.g. to feed post-processing tools. At the end of the experiment execution, the 5G EVE portal makes available a validation report with the analysis of the results for each test case, highlighting the KPIs values against the target thresholds.

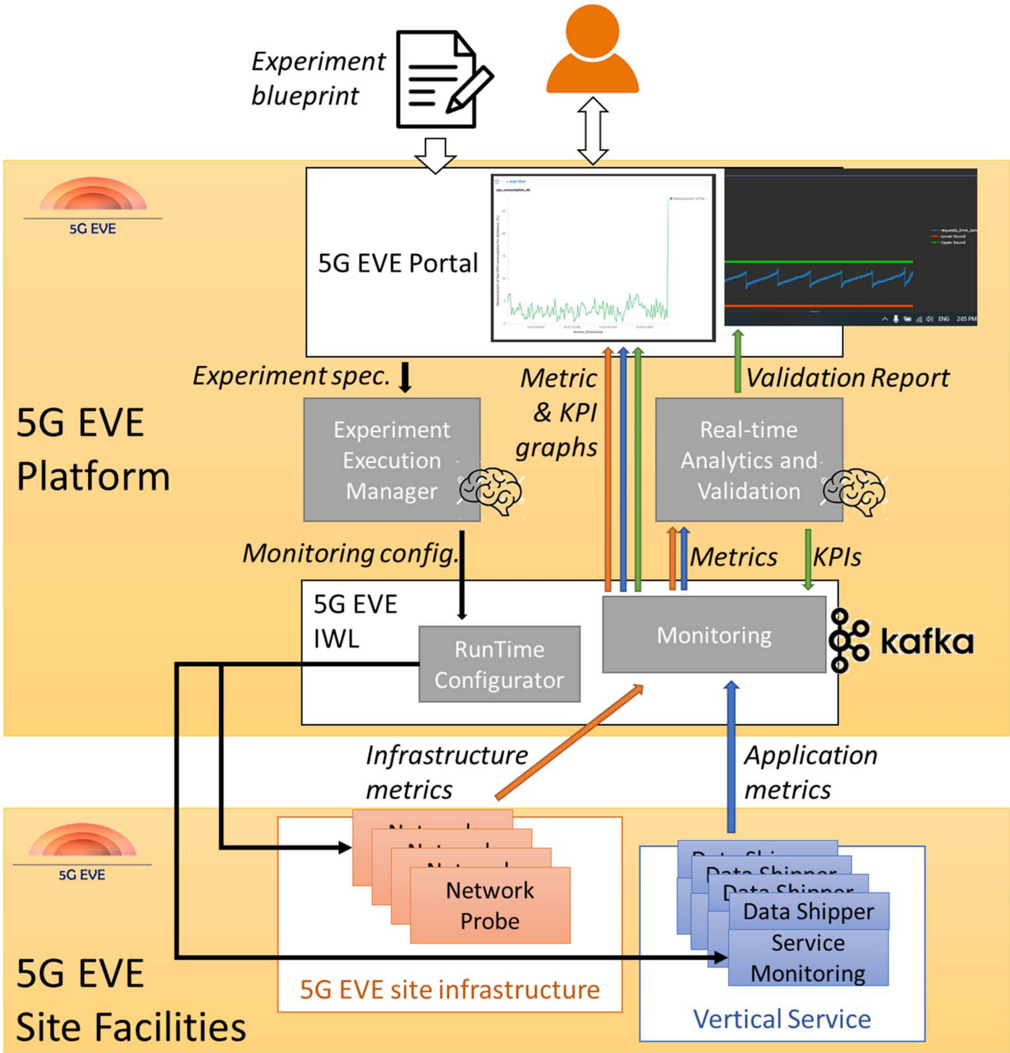

Figure A.5: 5G EVE Platform - procedures for metrics and KPI collection, validation and visualization

Figure A.5 shows the 5G EVE Platform internal procedures to collect, process and evaluate the metrics required to validate a vertical experiment. The 5G EVE portal processes the experiment blueprint retrieving the specification of the metrics and the KPIs and setting the related topics at the monitoring platform. The Kafka brokers at the monitoring platform, in the Data Collection Manager component, are thus ready to receive and exchange all the data related to the experiment on the configured set of topics. When triggering the execution of the experiment, the 5G EVE portal interacts with the Experiment Execution Manager (EEM) providing the experiment specification, including its monitoring topics. As first step of the execution, the EEM coordinates the configuration of the experiment elements. For the monitoring part, it interacts with the Run-Time Configuration (RTC) at the Inter-Working Framework providing the details of the monitoring configuration, like the topics, the list of infrastructure metrics, the IP addresses of the VNFs that run the data shippers for the application metrics, the scripts to configure these data shippers, etc. The RTC uses this information to configure the probes and the data shippers on the target functions. It should be noted that the list of network probes available on each facility and their capabilities is retrieved automatically from the site inventory at the Inter-Working Framework. In parallel, the Real-Time Analytics and Validation (RAV) component is configured with the list of topics, metrics, KPIs and thresholds, in order to activate the procedures for the automated analysis of the experiment results.

Once the data shippers have been properly configured and activated, the experiment is started and the metrics are continuously sent to the Kafka broker using the right topics. The 5G EVE monitoring platform relies on a hierarchical Kafka deployment, where monitoring data are initially collected on local Kafka brokers available in each site and replicated on a centralized one at the Inter-Working Framework. The RAV consumes the metrics from the centralized Kafka broker, on the basis of the per-topic subscriptions performed at the configuration phase. The metrics are processed and translated into KPIs that are pushed back into the Kafka broker. Both metrics and KPIs are stored in the Data Collection Storage of the 5G EVE monitoring platform and they can be visualized through graphs in the 5G EVE portal. At the end of the experiment, the RAV elaborates the entire set of KPIs, evaluating them according to the thresholds defined in the experiment blueprint. The results are summarized in a validation report, which can be visualized through the 5G EVE portal.

A.3 5G VINNI Platform

The 5G-VINNI Facility consists of multiple, inter-connected sites, each of which supports demonstration of a range of KPIs, using specific access technologies and end-user equipment types. Supported technologies include 5G-New Radio (NR) in sub-6GHz bands, 5G-NR in mm-wave bands, Low Power Wide Area (LPWA) networks and satellite.

A 5G-VINNI Facility is the deployment of the 5G-VINNI architecture in one administrative domain (e.g. one operator). The 5G-VINNI Facility-sites are classified into two different types:

- Main Facility-sites: E2E 5G-VINNI Facility that offers services with well-defined Service Level Agreements.
- Experimentation Facility-sites: 5G-VINNI sites that provide environments for advanced focused experimentation and testing possibilities on elements and combinations of elements of the E2E model.

The 5G-VINNI Facility-sites are illustrated in Figure A.6 with the Main Facility-sites (Norway, UK, Spain, Greece) and the Experimentation Facility-sites (Portugal, Germany/Munich, Germany/Berlin, Luxembourg).

Figure A.6: The 5G-VINNI Facility

Table A.1 lists the technical services offered by 5G-VINNI at the different Facility sites. Technical services here refer to resource facing services offered by the 5G-VINNI Facility; customer facing services are agreed and evolved according to the requirements of the verticals.

Table A.1: Technical services from 5G-VINNI which can be exposed to verticals

No.	Technical Services	Norway	UK	Spain	Greece	Portugal	Germany (Berlin)	Germany (Munich)	Luxemburg
1	eMBB slice	YES	YES	YES	YES	YES	YES	YES	YES
2	URLLC slice	YES	YES	YES	YES	YES	YES	YES	NO
3	mMTC slice	YES	YES	YES	YES	YES	YES	NO	YES
4	Autonomous core in the edge / Self-contained network	YES	NO	NO	NO	NO	YES	NO	YES
5	Fixed wireless access	YES	YES	YES	YES	NO	NO	NO	NO
6	Firewalling (Layer4-7)	YES	YES	YES	YES	YES	NO	NO	NO
7	Flexible backhaul for redundancy	YES	NO	NO	NO	NO	YES	NO	YES
8	Interconnection with Public cloud	YES	NO	NO	YES	YES	NO	NO	NO
9	Data fabric service involving correlation, aggregation and analytics	YES	YES	YES	NO	YES	NO	NO	NO
10	Test and KPI validation	YES	YES	YES	YES	YES	NO	NO	NO
11	3rd party VNF hosting	YES	YES	YES	YES	YES	NO	NO	NO
12	Edge cloud	YES	NO	YES	YES	YES	NO	YES	NO
13	Interconnection with other 5G-VINNI Facility-sites	YES	YES	YES	YES	YES	YES	NO	YES
14	Interconnection with non-5G-VINNI Facility-sites (to be offered based on demand)	YES	YES	YES	YES	YES	NO	NO	NO
15	Individual device connectivity (both eMBB and IoT) to 5G-VINNI Facility via default slice	YES	YES	YES	YES	YES	NO	NO	NO

Further information about the 5G-VINNI Facility can be found [i.61].

A.4 The 5G-VINNI Testing-as-a-Service system

5G-VINNI offers a testing infrastructure that is able to verify and validate the performance of the 5G-VINNI facility in terms of the 5G PPP KPIs. The testing infrastructure allows vertical customers to use the facility with a Testing-as-a-Service (TaaS) model, enabling the execution of dedicated campaigns with reduced effort. Open APIs and SDKs enable customers to integrate their own technologies within the framework.

Verticals' experiments can be performed through a user-friendly Testing Portal where tests can be composed, configured and executed. The results can be either visualized and analysed by tools offered in the Testing Portal, or be stored in an external database and processed with the vertical customers' own analysis and visualization tools.

The 5G-VINNI TaaS system is an automation and interfacing layer that allows connecting all the test and measurement tools needed for performing tests and experiments on the 5G-VINNI network. The automation allows to abstract the complexity with a series of either standard or custom Test Cases.

TaaS makes use of state-of-the-art products coming from partners' portfolios, prototypes, and open source components to offer vertical customers a tool for verifying network and application performance. Figure A.7 summarizes different test types and tool types that TaaS supports.

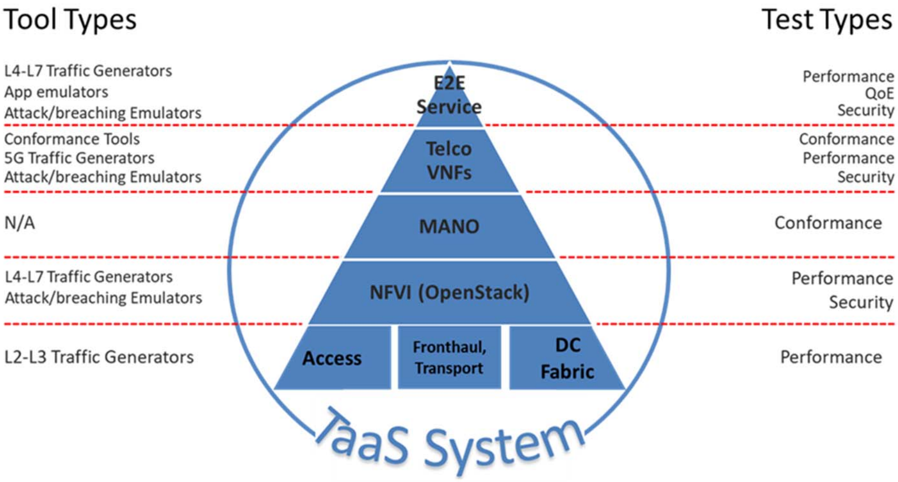

Figure A.7: 5G-VINNI TaaS system overview from a tooling perspective (from [i.61])

Verticals' own tools can be integrated into the TaaS system through plugins. These tools can then be configured and executed through TaaS in the same way as the tools offered by 5G-VINNI.

TaaS is a cloud based system with a general architecture shown in Figure A.8.

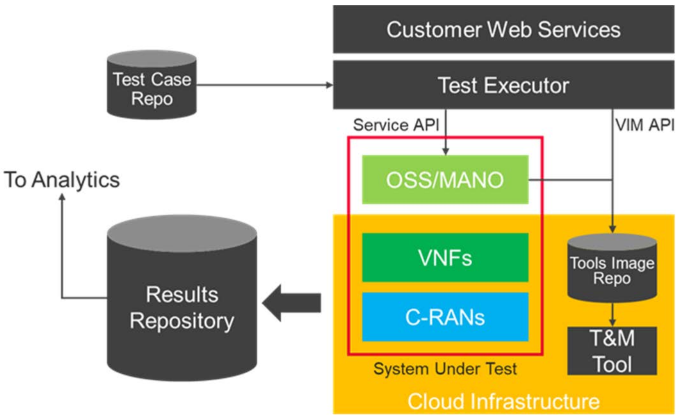

Figure A.8: TaaS general architecture

The **Customer Web Service** allows human users to define, create, and execute test campaigns. An API that can be contacted by other applications (e.g. CI/CD pipelines) for consuming testing services are also offered. The available test scripts for configuring and executing tests are stored and managed in the **Test Case Repository**. The **Test Executor** coordinates the different tools needed and performs the tests. The **VIM** allows the Test Executor to deploy the needed infrastructure and tools to perform the tests. The **Test and Measurement (T&M)** tools are used to insert traffic into and probe the 5G infrastructure. The results generated by the tools are stored in the **Results Repository**.

A.5 5GENESIS Platform

The 5GENESIS platform implements and verifies evolutions of the 5G standard (from 3GPP Rel 15 onwards), via iterative integration and testing procedures. Heterogeneous physical and virtual network elements are unified under a common coordination and openness framework that is exposed to experimenters/vertical industries and enables end-to-end slicing and experiment automation. More precisely, the 5GENESIS platform abides by the architectural principles defined in the 5GENESIS project, namely, modular set up, openness and automation, and experimentation process formalization:

- **Modular set up:** Three well interfaced components are defined, namely the *infrastructure*, the *network management & orchestration*, and the *experiment coordination*. The interaction among the components is well defined, to facilitate any replacement/update/change in one of the components without affecting the others. A web portal with the experimenters/vertical industries has been developed as well, to enable:

  i) the potential for use case- specific configuration; and

  ii) visualization of the performance results and, also, analytics on the data collected during the experiments.

- **Openness and automation:** Open source software components (i.e. the OAI) have been adopted to implement the mobile network functionality (RAN and core network functions). Also, all the management and coordination layer features are openly released by 5GENESIS project under the term *Open 5GENESIS suite (https://github.com/5genesis)*. The strategic selection of automation tools, such as the open TAP (https://www.opentap.io/), provides experimenters with guarantees, regarding the accuracy and reliability of the measurements (multiple experimentation scenarios can be repeated under the same conditions).

- **Experimentation process formalization:** In addition to the developments and the integrations conducted, measures to facilitate the experimenters during the procedure of running an experiment have been taken. To this end, the information required for running an experiment has been formulated. The formulation led to a set of useful templates available to the experimenter, as well as a pool of available predefined tests. In addition, a set of vertical specific software pieces and configurations have been produced to enable an easy and fast experimentation process for default scenarios, and to provide educating examples to vertical industries.

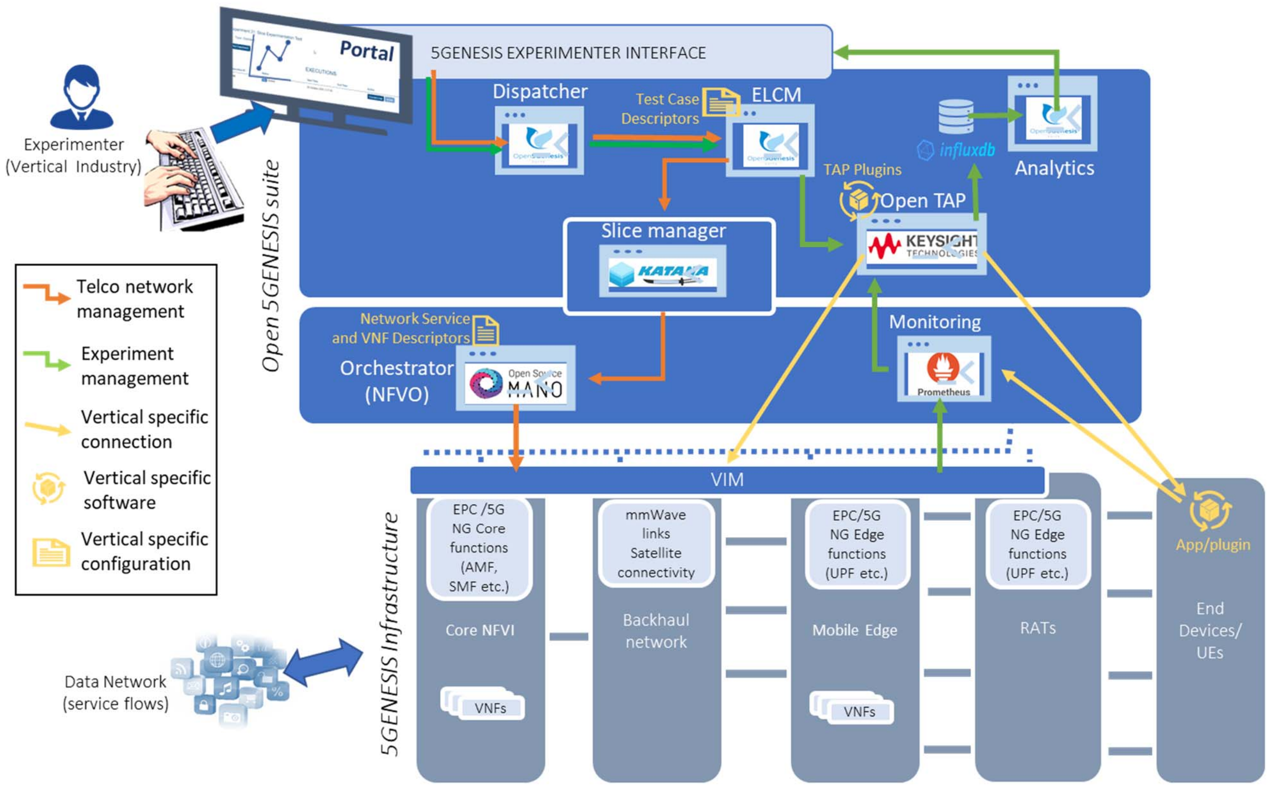

Figure A.9: 5GENESIS Platform - major architectural components

The infrastructure component of the 5GENESIS platform engages a wide diversity of technologies and chain innovations that span over all network domains, achieving full-stack coverage of the 5G landscape. Five end-to-end network developments have been realized, in the sense that end-user devices, as well as the radio access, core, backhaul and transport domains are part of the service provisioning chain. Each one of the five developments provides different capabilities and integrates advances required from different 5G use cases. More precisely, the 5GENESIS Athens infrastructure focuses on validating capabilities of the edge in small cell coverage; the 5GENESIS Berlin infrastructure focuses on large deployment with point-to-point mmWave links in the transport part of the network, the 5GENESIS Malaga infrastructure focuses on mission critical services in urban areas supported by the edge; the 5GENESIS Surrey infrastructure focuses on dense user deployments; and the 5GENESIS Limassol infrastructure focuses on the integration of Satellite communications in the service provisioning chain.

On top of the infrastructure, two management and control flows are supported through a web interface/portal. For both the flows the inputs from the experimenter/vertical are authorized and validated in an entity called *dispatcher*. Then the choices/inputs are passed to the *Experiment LifeCycle Manager (ELCM)*. The first flow refers to the configurations needed for the virtualized part of the MNO functions, i.e. the network functions of the 3GPP 5G SBA, such as the AMF, SMF, UPF etc. The key entity in this flow is the slice manager. The slice manager realizes the functionality of Network Slice Management Function (NSMF), as defined by 3GPP, and utilizes The Generic Network Slice Template (GST) v2.0 as provided by GSMA. A set of NESTs (NEtwork Slice Types) i.e. GSTs filled with values, is also available for experimentation. The second management and control flow that originates at the portal refers to the experimentation process. The main entity in the path of this flow is the *automation tool (openTAP)*. The experimenter/vertical selects the test cases to be executed and the ELCM commands the appropriate plugins through the automation tool. Based on the vertical industry, mobile applications/probes/plugins at the end device could be added as well. In case that vertical specific software should be deployed as a VNF in the core and edge nodes, it can be done through openTAP.

KPIs can be monitored at any node of the end-to-end service provisioning chain, i.e. at any user plane node of the infrastructure. The collected KPIs define a pool of measurements that feed the analytics tool of the platform. The analytics tool provides: statistical analysis, results visualization, variables correlation and dependency check, as well as performance projection based on Machine Learning mechanisms.

Overall, an experimenter/vertical industry, interested in the utilization of the 5GENESIS platform for 5G-app validation should: select the infrastructure that best fits to the services/app that it offers, fill in (together with the infrastructure leader) an experiment descriptor, develop (or request the development of) the plugins that the experiment requires, execute the tests and collect the results and the analysis report.

Annex B:
EC FIRE programs

Within the FP7-ICT - Specific Programme "Cooperation": Information and communication technologies the European Commission (EC) issued the ICT-2011.1.6 - Future Internet Research and Experimentation (FIRE). Furthermore, within the H2020-EU.2.1.1. - INDUSTRIAL LEADERSHIP - Leadership in enabling and industrial technologies - Information and Communication Technologies (ICT) the European Commission issues the call ICT-13-2016 - Future Internet Experimentation - Building a European experimental Infrastructure.

The aim of both FIRE and FIRE+ programs was providing support for building infrastructures for the design and deployment of products, applications, and services on the Future Internet. Such infrastructure were made available to experiments of any size, complexity, or networking technology. Experimenters were capable of running experiments under controlled and replicable conditions, according to specific requirements by accessing real or virtual equipment, services, systems and tools on demand, seamlessly and regardless of their geographical location.

Such infrastructures were aimed to achieve the following goals:

- experimental capability at European level that covers a variety of networking technology areas and allows tens of experiments to be run on top of them each year;
- potential to experiment without the constraints of the physical location or access to a specific experimental facility;
- reduction of the time to experiment by allowing a larger set of experiments to take place on reliable and benchmarked infrastructure that can evolve and be re-configured;
- response to the needs of individual, small and medium experimenters without access to experimental facilities or environments;
- support of trials driven by vertical application areas with a good mix of supply and demand stakeholders;
- contribution to the sustainability model of experimental facilities;
- contribution to standardization and interoperability of experimental facilities.

Some successful examples of FIRE+ projects are, to name a few, Fed4FIRE+ and TRIAGLE.

In Fed4FIRE+ (https://www.fed4fire.eu/) a federation of worldwide of Next Generation Internet (NGI) testbeds, which provide open, accessible and reliable facilities supporting a wide variety of different research and innovation communities and initiatives in Europe, including the 5G PPP projects and initiatives. Fed4FIRE+ offered external experimenters the possibility of running experiments on several facilities through a portal and by reserving slices of resources through a middleware (jFed).

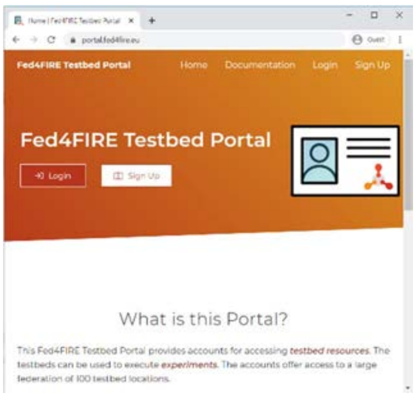

Figure B.1: Fed4FIRE+ Portal

In TRIANGLE (https://www.triangle-project.eu/) a framework was built to help app developers and device manufacturers in the evolving 5G sector to test and benchmark new mobile applications in Europe utilizing existing and extended FIRE testbeds. This framework offered the possibility of evaluating Quality of Experience and enable certification for new mobile applications and devices.

The project identified reference deployment scenarios, defined new KPIs (Key Performance Indicators) and QoE metrics, developed new testing methodologies and tools, and designed a complete evaluation scheme. The framework, methods and tools developed during the project focused on providing the mechanisms to incorporate new wireless technologies and topologies envisaged in 5G and contribute to the new ecosystem.

Annex C: SLA requirements on "Predictive QoS" for Automotive / V2X applications and on "Time Sensitive KPIs"" for Industry 4.0

C.0 Background

This annex addresses the topic on "Inter-Domain and Inter Operator options for E2E Network Slice Autonomic Service and Security Assurance for Verticals' applications' SLA requirements on "Predictive QoS"" for Automotive / V2X applications and on "Time Sensitive KPIs"" for Industry 4.0", extracted from clause 7.5 of the ETSI TC INT PoC Whitepaper #4 [i.73] (https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals).

C.1 Use Cases of Federated AMC Knowledge Planes (Inter-Domain and Inter Operator option) for E2E Network Slice Autonomic Service Assurance

From Business perspective, those use cases use Dynamic Network Slice Management through Autonomic Service Assurance process powered by AI capabilities as a means for CSPs to fulfil the requirements of Vertical Industries (e.g. Industry 4.0, Automotive, Bank, Insurance, Smart Cities, etc.) while deploying and operating a single 5G Network either partitioned into Multiple Domains or federating Multiple Operators' 5G Networks when required.

Examples of Verticals' applications SLA requirements are "Predictive QoS"" for Automotive/V2X applications and "Time Sensitive KPIs"" for Industry 4.0 that should be Timely, Dynamically and on-fly (re)- negotiated against 5G Network supported capabilities. This is similar to the traditional process of SLA Management (Negotiation, Operations / Execution / Enforcement, Report) as documented in the TMForum IG1127 [i.74] (reference "End-to-end Virtualization Management: Impact on E2E Service Assurance and SLA Management for Hybrid Networks". Indeed, both the "contractual" SLA between two parties (a Customer and a Provider) and the technical-level QoS characterization use one common template: the SLS (Service Level Specification), which is a detailed list of Metrics ("SLS Parameters") and associated SLS Thresholds (to trigger SLS Consequences when expected service levels are not met). SLS Thresholds are informally also known as SLO (Service Level Objective). Network Slice as a "product" provided and consumed is associated with SLAs hence follows this SA Management process (TMForum IG1127 [i.74]) while using GSMA Generic Network Slice Template Version 1.0, 23 May 2019 and 3GPP 5G terminology specification 3GPP TS 28.531, [i.71].

This aspect is reflected at the left hand side (green part) of the three Figures C.1, C.2 and C.3. Indeed the negotiated SLA as a first step to agreed SLA between Network Slice Producer and the vertical aims at matching Verticals Industries SLA requirements (a) with 5G Network Slice Producer (5G Operator(s)) capabilities (a). The result leads to the agreed result (b). This is achieved by successively instantiating (populating different fields with values or range of values) from (c) GSMA GST (Generic Slice Template), (d) GSMA NEST (Network Slice Type), (e) 3GPP NST (Network Slice Template) [3GPP TS 28.531 [i.71]] from whichcan be derived multiple Network Slice Instances. Phase (f) corresponds to injection and Autonomic orchestration, self-monitoring of the SLA in the 5G Network(s) Autonomic-capable. This NEST corresponds to "5G Design Template" defined five years ago in the description of the high level design principle of the ETSI TC INT 5G PoC ecosystems and associated actors / roles relationships and interactions as depicted in "5G PoC White Paper N°3, Figure 12" https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals. Its translation onto required resources by the E2E Service Orchestrator corresponds per Network Slice Instance corresponds to the NST.

Phase (g) corresponds to the real-time reporting of the status of consumed Network Slice Instances by making available to the involved Verticals (Networks Slice Instances Consumers) customized Real-Time (RT) Dashboards on consumed 5G Network Slices per Network Slice Instance, per Application, per Device according to Vertical Industries request.

This continuous adjustment and updates of the SLAs is realized through GANA - AMC Autonomic (Multi-layer / Nested Control-Loops) Federated Framework. White paper N° 3 on "5G PoC Programmable Traffic Monitoring for Network Slices Service Assurance" [i.75] https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals is a detailed description, implementation and demonstration of this use case.

Besides this described Dynamic SLA Management process Involving Network Slice Provider and Vertical Industries as Slices Consumers, it also involves internal stakeholders (Inter-Domain model figure C.1 and figure C.2 where internal SLA named OLA (Operational Level Agreement) at touch points between 5G Domains (h) is managed. In the Inter Operator model (figure C.3), there is also SLAs at the touch points (i).

It may happen that 3rd Parties (e.g. SLA Broker) play certain roles between Provider and Consumer by offering SLA Management services or certain SLA responsibilities might be delegated to them by the one, or more main parties. The SLA Broker is connected through doted lines (j) as depicted in Figure C.3. Those four doted lines represent the respective potential SLA contracts (from negotiation to execution then reporting).

In terms of design principle illustrated by the diagrams in the Figures C.1, C.2 and C.3 those use cases follow the same approach as the ones uses in sections 7.1, 7.2, 7.3, 7.4 of the ETSI TC INT PoC Whitepaper #4 [i.73] (https://intwiki.etsi.org/index.php?title=Accepted_PoC_proposals). (CSP (s)' Autonomic Production Network and related Training & Testing Environment for GANA -AMC DEs / AI Models and external stakeholders such as GANA-AMC DEs Marketplace, Auditor / Regulator, etc.

Two main options are considered:

- Option-A Horizontal Federation in Inter-Domain model (Single Organization)
- Option-B Vertical Federation in Inter-Domain model (Single Organization)
- Option-B Vertical Federation in Inter-Operator model (Multi Organization)

What is common to those options?

F-MBTS: A federation translation function (F-MBTS) may be required if data models and communication methods for federations employed by the two or more domains are different.

AMC-MBTS: It is a translation function placed between the Network Layer and the Knowledge Plane.

ONIX: As a Knowledge Base that can acts as a Real-Time Inventory: in the Inter-Domain model (Single Organization) a unique and shared ONIX may be the option. In the Inter-Operator model (Multi Organization) a fragmented (dedicated) ONIX is the appropriate approach as each Operator own it Knowledge Data base and Data are structured in a specific format that could be different form the one of the other Operators engaged in this federation.

C.2 Description of possible options

C.2.1 Option-A (Horizontal Federation) in Inter-Domain model (a single Organization)

It is the option which the GANA Knowledge Plane (KP) Platforms for the specific network segments (e.g. 5G RAN, X-Haul, 5G Core Network) federate horizontally with each other without the need for an overlay umbrella Hierarchical GANA Knowledge Plane (KP) Platform.

In such an option there is a need for an Interworking / Coordination Reference Point for E2E Federation of Knowledge Planes (e.g. 5G RAN - KP, Xhaul- KP, 5G Core Network - KP.

Figure C.1 depicts this option A.

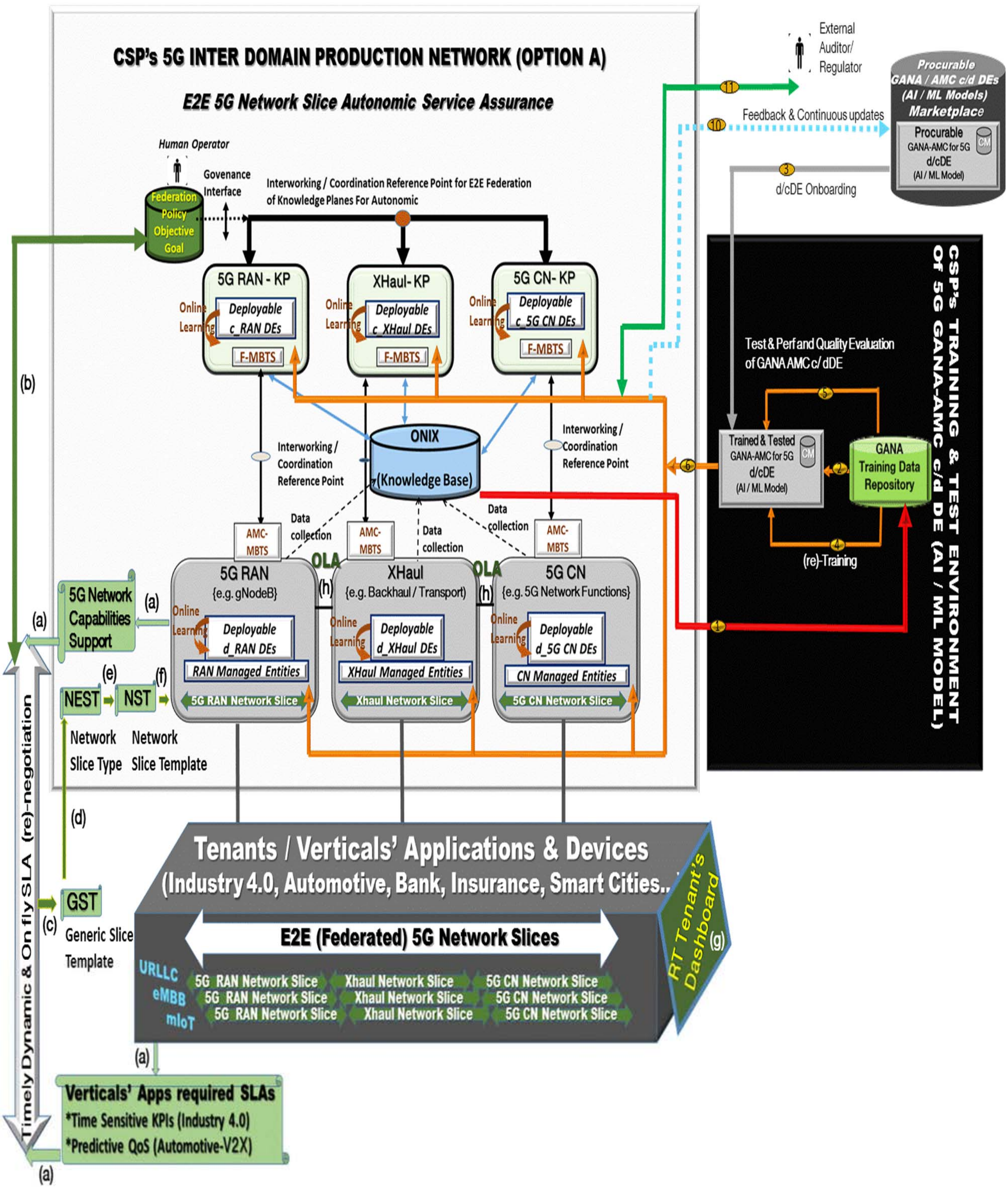

Figure C.1: Option-A (Horizontal Federation) in Inter-Domain model (a single Organization)

C.2.2 Option-B (Hierarchical /Vertical Federation) in Inter-Domain model (a single Organization)

It is the option by which the GANA Knowledge Plane (KP) Platforms for the specific network segments (e.g. 5G RAN, X-Haul, 5G Core Network) federate vertically through an overlay umbrella Hierarchical / Vertical GANA Knowledge Plane (KP) Platform or "Inter-Domain Knowledge Plane " that receives information from the lower level KPs (e.g. 5G RAN - KP, Xhaul- KP, 5G Core Network - KP) and coordinates the lower level KPs. Figure C.2 depicts this option B.

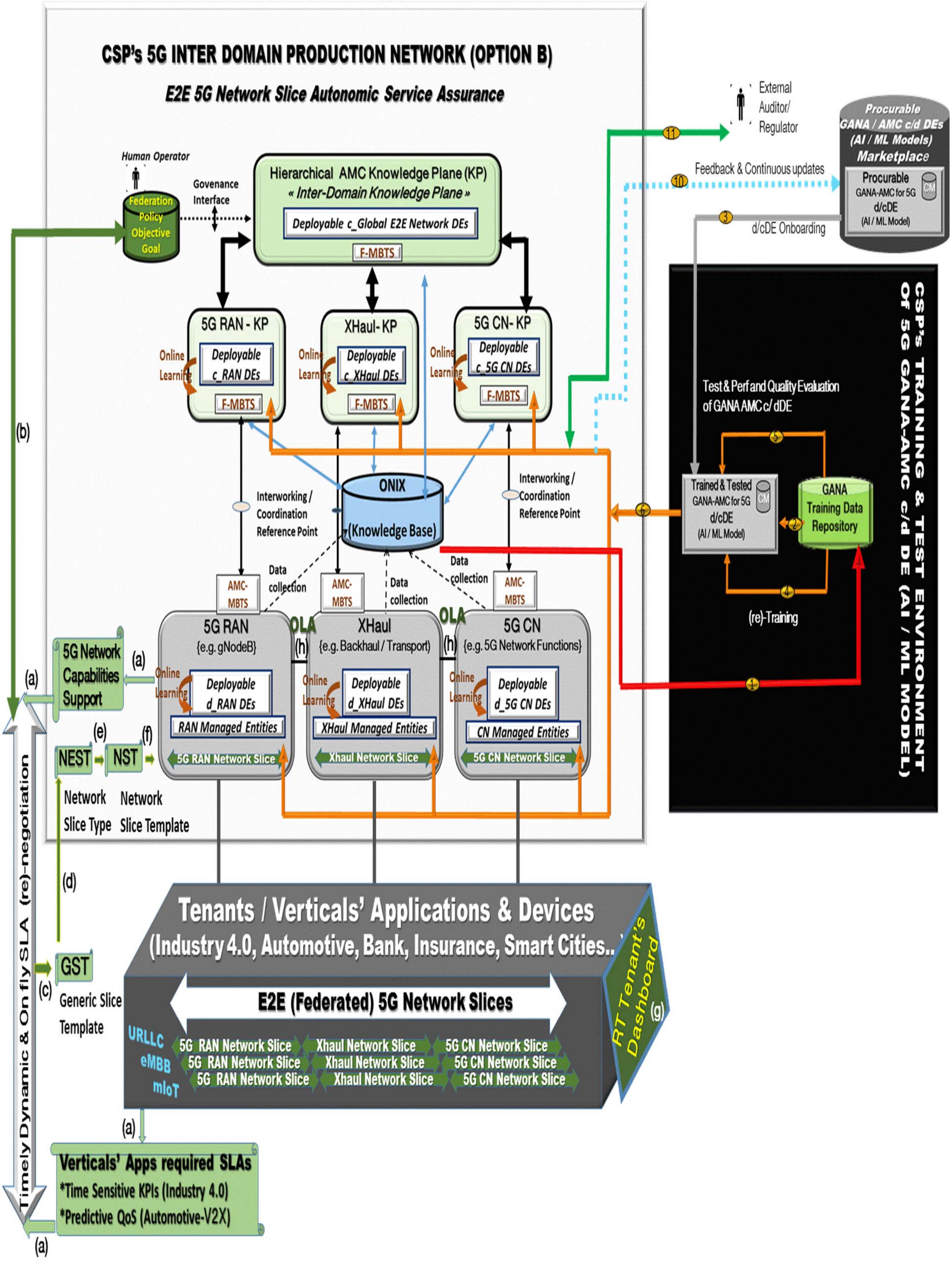

Figure C.2: Option-B (Hierarchical /Vertical Federation) in Inter-Domain model (a single Organization)

C.2.3 Option-B (Hierarchical /Vertical Federation) in Inter-Operator model (Multi Organization)

It the option by which the GANA Knowledge Plane (KP) Platforms for collaborating Operators' Network (e.g. Operator #A 5G Network, Operator #B 5G Network, Operator #C 5G Network) federate vertically through an overlay umbrella Hierarchical / Vertical GANA Knowledge Plane (KP) Platform or "Inter-Operator Knowledge Plane" that receives information from the lower level KPs (e.g. Operator #A 5G Network - KP, Operator #B 5G Network - KP, Operator #C 5G Network - KP) and coordinates the lower level KPs. Figure C.3 depicts this option B.

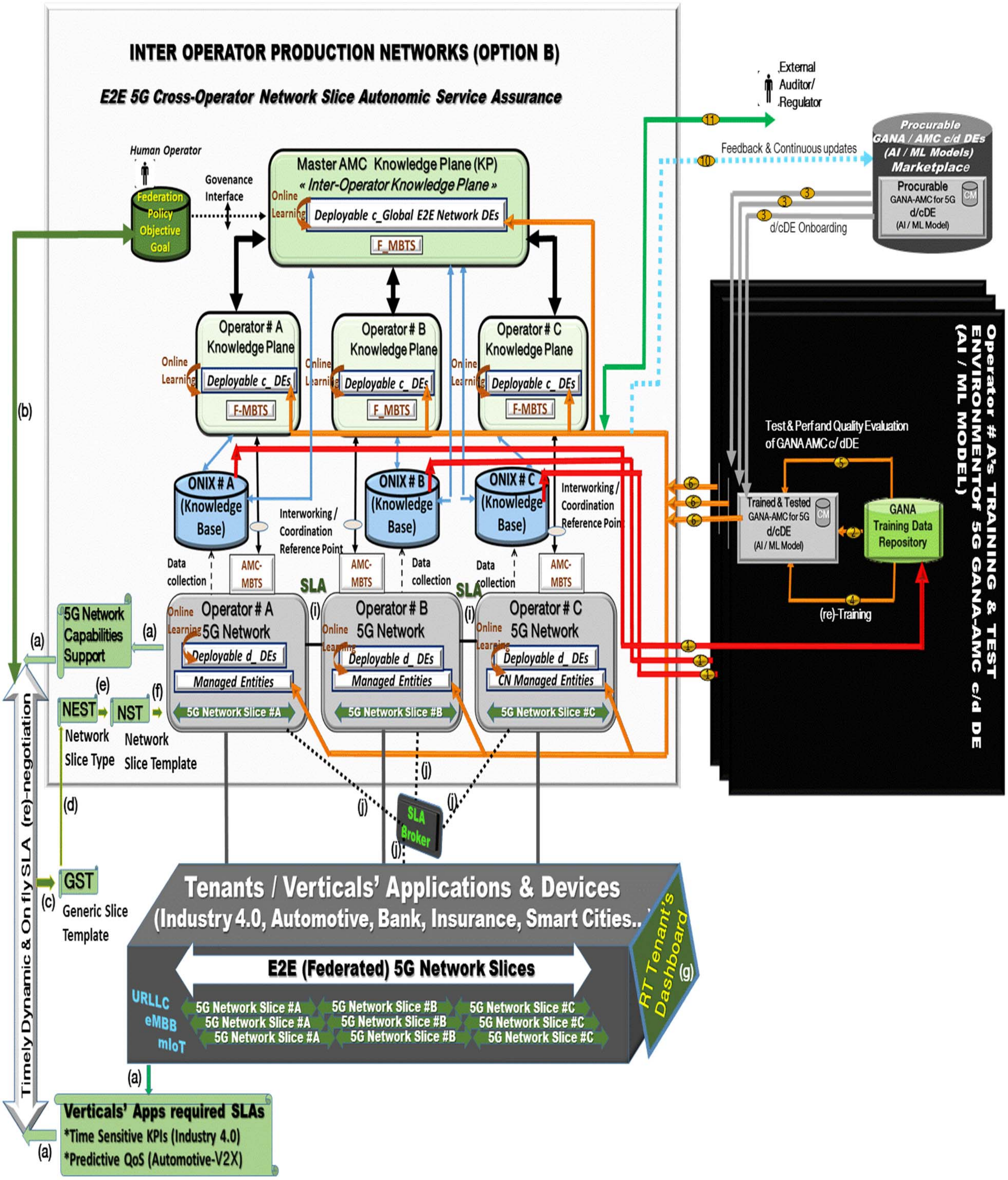

Figure C.3: Option-B (Hierarchical /Vertical Federation) in Inter-Operator model (Multi Organization)

History

Document history		
V1.1.1	May 2022	Publication