5Genesis

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

[H2020 - Grant Agreement No. 815178]

Deliverable D6.2

Trials and experimentation (cycle 2)

Editor G. Xilouris, M. Christopoulou (NCSRD)

Contributors NCSR Demokritos (NCSRD), Fraunhofer (FhG), Innovations for High Performance microelectronics (IHP), Humboldt University of Berlin (HUB), Karlstads Universitet (KAU), Simula Research Laboratory (SRL), RUNEL (REL), EURECOM (ECM), Infolysis (INF), Intel (INT), University of Surrey (UNIS), Fogus (FOG), Nemergent (NEM), , University of Malaga (UMA), Cosmote (COS), Universitat Politecnica de Valencia (UPV), Atos (ATOS)

Version 1.0

Date June 30th, 2020

Distribution PUBLIC (PU)

List of Authors

Fraunhofer	FOKUS-Fraunhofer Gesellschaft e.V., Institute for Open Communication Systems
M. Emmelmann, M. Monachesi, R. Shrestha, F. Eichhorn	
IHP	Innovations for High Performance microelectronics/ *Leibniz-Institut für innovative Mikroelektronik*
J. Gutiérrez, M. Ehrig, N. Maletic, E. Grass	
KAU	Karlstads Universitet
A. Brunstrom, M. Rajiullah	
SRL	Simula Research Laboratory AS
G. Caso, C. Griwodz	
REL	RUNEL NGMT LTD
I. Koffmann	
ECM	EURECOM
P. Matzakos, F. Kaltenberger	
INF	INFOLYSiS
V. Koumaras, C. Sakkas, G. Theodoropoulos, A. Papaioannou	
INT	INTEL
V. Frascolla	
UNIS	University of Surrey
S.Vahid	
FOG	Fogus
D. Tsolkas	
NEM	NEMERGENT SOLUTIONS SL
E. Alonso	
NCSRD	NATIONAL CENTER FOR SCIENTIFIC RESEARCH “DEMOKRITOS”
G. Xilouris, M. Christopoulou, H. Koumaras, T. Sarlas, T. Anagnostopoulos	
UMA	UNIVERSITY OF MALAGA
A. Díaz, I. Gonzálezi, B. García, P. Merino	
COS	COSMOTE KINITES TILEPIKOINONIES AE
I. Mesogiti, F. Setaki, E. Theodoropoulou	

Disclaimer

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

Copyright

The 5GENESIS Consortium consists of:

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

Version History

Rev. N	Description	Author	Date
1.0	Release of D6.2	G. Xilouris (NCSRD)	30/07/20

LIST OF ACRONYMS

Acronym	Meaning
5G	5-th Generation of cellular mobile communications
5G NR	5G New Radio
5G-PPP	5G Public-Private Partnership
AMQP	Advanced Message Queuing Protocol
API	Application Programming Interface
ARM	Advanced RISC Machine
CN	Core Network
COAP	Constrained Application Protocol
COTS	Commercial-Off-The-Self
DL	Downlink
DRX	Discontinuous Reception
DUT	Device Under Test
DRAN	Distributed Radio Access Network
DWDM	Dense Wavelength Division Multiplexing
E2E	End-to-End
ELCM	Experiment Life Cycle Manager
eMBB	Enhanced Mobile Broadband - 5G Generic Service
EMS	Element Management System
EPC	Evolved Packet Core
E-UTRAN	Evolved Terrestrial Radio Access Network
FCAPS	Fault, Configuration, Accounting, Performance and Security
FPGA	Field Programmable Gate Array
GDPR	General Data Protection Regulation
HEVC	High Efficiency Video Coding
IaaS	Infrastructure as a Service
IED	Intelligent Electronic Devices
ING	Intelligent Network Gateway
IUG	Intelligent User Gateway
KPI	Key Performance Indicator
LOS	Line of Sight
LTE	Long Term Evolution
MANO	Management & Orchestration

MIMO	Multiple Input Multiple Output
MCPTT	Mission critical push-to-talk
MCS	Mission critical services
MME	Mobility Management Entity
mmWave	Millimetre Wave
NB-IoT	Narrow Band – Internet of Things
NFVO	Network Function Virtualization Orchestrator
NLOS	Non Line of Sight
NMS	Network Management System
NSA	Non-Stand-Alone
OAI	Over the Air Integration
OSS	Operational Support Services
PFCP	Packet Forwarding Control Protocol
PLMN	Public Land Mobile Network
PSM	Power Saving Mode
PTMP	Point-to-Multi-Point
QoE	Quality of Experience
QoS	Quality of Service
PTMP	point-to-multipoint
RAT	Radio Access Technology
REST	Representational State Transfer
RRH	Remote Radio Head
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RTP	Real-time protocol
RTSP	Real-time Streaming Protocol
SA	Stand-Alone
SDK	Software Development Kit
SDN	Software Defined Networking
SINR	Signal to Interference plus Noise Ratio
SUT	System Under Test
TaaS	Testing as a Service
TAP	Test Automation Platform
TTN	The Things Network
UL	Uplink
UMA	University of Málaga

VIM	Virtual Infrastructure Manager
VM	Virtual Machine
VNF	Virtual Network Function
VNFM	Virtual Network Function Manager
VPN	Virtual Private Network
VR	Virtual Reality
WIM	WAN Infrastructure Manager
WLAN	Wireless Local Area Network
WP	Work Package

Executive Summary

This deliverable presents the second cycle of trials and experimentation activities executed over 5GENESIS facilities. The document is the continuation of deliverable D6.1, in the sense that it captures tests carried out over the evolved infrastructures hosting 5GENESIS facilities following the methodology defined in D6.1. In this document 8 main KPIs and 4 application specific validation trials achieved, under 123 experiments that performed in total. The tests focus more on i) the evolved 5G infrastructure deployments that includes radio and core elements in non-standalone (NSA) deployment configurations based on commercial and open implementations, and ii) the use of Open 5GENESIS Suite for the execution of the tests.

In this context the structure of this document is platform centric, hence it allows each platform to specify independently the group of executed tests and validations executed and the results presented and commented. However, all platforms agree on the following principles:

- Test and results discussed in this document should be accompanied by the detailed Test Case description according to 5GENESIS template specified in deliverable D2.3 [1].
- Throughput and Round-Trip-Time (RTT) Key Performance Indicators (KPIs) should be validated in all platforms although the results are not mandatory to be comparable.
- The test procedure should be carried out through the portal / Experiment Lifecycle Manager (ELCM) components of the Rel. A of Open 5GENESIS Suite.
- The result analysis should be carried out through the analytics tools provided by the Open 5GENESIS Suite analytics framework.

In most of the cases all the previous clauses were applied, however some specific tests either due to the complexity of the scenario, the advanced required functionality or the on-going status of the deployment, were executed manually.

In brief, five 5GENESIS facilities' tests and results are discussed in this document:

- Málaga Facility - includes baseline tests for throughput and RTT KPIs, as well as MCPTT, RunEl 5G Radio Access Network (RAN) solution physical layer latency and content delivery streaming service.
- Athens Facility - includes baseline tests for throughput and RTT KPIs, Service Creation Time of a 5G connectivity service, one-way delay and RTT under different load scenarios.
- Limassol Facility - includes baseline tests for throughput and RTT KPIs as well as Service Creation Time of 5G component deployment.
- Surrey Facility - includes baseline tests for throughput and RTT KPIs, NB-IoT coverage, IoT application specific latency.
- Berlin Facility - include baseline tests for throughput and RTT KPIs, 360° live video streaming QoE, RAN coverage and UE density.

The main part of the document contains a basic presentation of the validated KPIs and measured metrics followed by commentary. The detailed test cases and result tables are available in the Annex of this document. It should be noted that the original Test Cases (the ones available in the annex of D6.1) have been refined and are delivered as a separate Testing

and Validation companion document. This document includes all the test cases templated (i.e. the KPI measured, the System Under Test (SUT) definition, the measurement process and tools) that have been used throughout D6.2 and D6.1. In addition, in a separate section the common 5GENESIS measurement statistical analysis methodology is summarised.

Overall, deliverable D6.2 presents the progress of 5GENESIS for the last year (2nd cycle) and the results already reveal the benefits from the adoption of 5G technology. In some cases, included in this document, the previous statement is also verified for preliminary vertical applications implementation.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

The aim of the 5GENESIS project is to evaluate and validate various 5G equipment and network deployments (such as those comprising the five 5GENESIS platforms), towards the achievement of the KPIs' targeted values with respect to those expected in commercial 5G network deployments. In addition, 5GENESIS will also deploy, validate and demonstrate various vertical cases on top of the aforementioned platforms.

The work in the field of KPI validation and performance evaluation is shared among the platforms but also work in complementary mode especially in some set of baseline KPIs (i.e. latency, throughput, service deployment time) which are common to all platforms and given the variety of the infrastructure implementation in each platform their comparison may have some added value.

This deliverable describes the trials and experimentation results from the second integration cycle of 5GENESIS. Upcoming versions of this deliverable will describe the trials and experimentation results from the third integration cycle (D6.3, M36). To better depict the progress conducted, it is expected that those documents will maintain similar structure as this deliverable.

1.1. Purpose of the document

These deliverable results are obtained by experimentation procedures that were conducted over the five 5GENESIS facilities where the Open 5GENESIS suite was integrated. In this context the interdependencies of this document are presented in Table 1-1.

Table 1-1 Document interdependencies

id	Document title	Relevance
D2.1 [2]	Requirements of the Facility	The document sets the ground for the first set of requirements related to supported features at the testbed for the facilitation of the Use Cases.
D2.2 [3]	5GENESIS Overall Facility Design and Specifications	The 5GENESIS facility architecture is defined in this document. The list of functional components to be deployed in each testbed is defined.
D2.3 [1]	Initial planning of tests and experimentation	Testing and experimentation specifications that influence the testbed definition, operation and maintenance are defined.
D3.1 [4]	Management and orchestration (Release A)	The document presents the MANO solutions that are integrated in the infrastructure. Interfaces and deployment options are also described.

D3.5 [5]	Monitoring and analytics	This document presents the methods and the framework for obtaining the statistical results for all the experiments in the deliverable
WP4 Del.	Athens D4.2 [6] Malaga D4.5 [7] Limassol D4.8 [8] Surrey D4.10 [9] Berlin D4.14 [10]	These documents describe the platform setup, capabilities and level of integration.
D6.1	Trials and Experimentation (cycle 1) [11]	This document is updated according to phase 2 evolution at the testbeds and coordination layer framework. The results were obtained using the D6.1 methodology and updated test descriptions and measurement procedures.

1.2. Structure of the document

The document is devoted to the presentation of the experimental results obtained in the second phase of 5GENESIS project, updating and/or complimenting the results of D6.1. The first part of the document (main document) is devoted to experiments and trials that were conducted in each 5GENESIS facility. Separate sections are devoted for each facility. The document concludes discussing the results across testbeds. The second part of the document is devoted to the detailed testing procedures and received results from each facility. Finally, the document is accompanied with an additional test companion (provided as a separate document) containing all the 5GENESIS Test Cases [12] used for the presented experimental results.

1.3. Target Audience

The primary target audience of this first WP6 deliverable encompasses industry and standardization stakeholders, allowing them to validate the 5G KPIs, based on the description of the test cases and the subsequent experimentation results from the first integration cycle, providing the joint evaluation of the results obtained from the experiments in the different platforms.

As the approach is based on industry best practices, this deliverable is best suited for industry stakeholders, although not limited to them.

Other stakeholders that can benefit from the document include:

- Standardisation organizations
  Where the test cases can form the basis of test suites.

- European Commission
  To evaluate the conduction and results of 5G experimentation.
- Academic and research stakeholders
  As basis for design decisions for 5G based frameworks and applications development.
- Non-experts interested in 5G opportunities
  To understand the capabilities and limitations of 5G technology.

2. METRICS AND TEST CASES

Test case IDs were defined in deliverable D6.1 [11] to refer to the addressed metrics. The test case IDs that are used in this document are further detailed in the companion document "5GENESIS TEST CASES v.1.0"[1] and are presented in Table 2-1. It should be noted that this table also contains test cases that were defined in D6.1 [11] but are not used in this document.

Calibration tests were part of deliverable D6.1 and are not included in this document. Essentially calibration tests are required prior to the initiation of any measurement or validation conducted on the platform and the responsible to run these tests is the platform owner/ operator. Any tester or experimenter should assume that the facility is always working as it supposed to be. In case of pre-test calibrations, e.g. calculation of buffering delay in order to be removed from the overall end-to-end (E2E) delay calculations, these calibrations are part of the test case description.

Table 2-1 Test Case and KPI mapping

KPI	Primary Metric	Test Case ID
Capacity	Throughput	TC_CAP_AreaTrafficCapacity
Density	# reg. users	TC_DEN_MaxRegisteredUE_BER_001
	# active users	TC_DEN_MaxActiveUE_BER
	# operations per sec	TC_DEN_MaxOpReqProcessed_BER
	time to register	TC_DEN_OpProcessingDelay_BER
	# reg. users	TC_DEN_MaxRegisteredUE_BER_002
Energy Efficiency	Energy Consumption	TC_ENE_RANEnergyEfficiencyAVG
	Energy Consumption	TC_ENE_RANEnergyEfficiencyMAX
	Energy Consumption	TC_ENE_UEEnergyEfficiency
	Energy Consumption	TC_ENE_NBIoT_SUR
Latency	one way latency	TC_LAT_e2eAppLayerLatency
	one way latency	TC_LAT_PHYLatency_MAL
	one way latency	TC_LAT_SmartGridControlMsgLatency_BER
	one way latency	TC_LAT_APPLayerLatency
Round Trip Time	RTT	TC_RTT_e2e
	RTT	TC_RTT_e2eBGTraffic
	RTT	TC_RTT_e2eRadioLinkQuality
Service Creation Time	time elapsed	TC_SCT_VMDeploymen_BER
	time elapsed	TC_SCT_5GConnSliceInstantiation
Throughput	data rate	TC_THR_Tcp
	data rate	TC_THR_Udp
Ubiquity/Coverage	packet loss	TC_UBI_RANCoverage
	packet loss	TC_UBI_BHCoverage

---

[1] https://github.com/5genesis/5genesis_test_cases/blob/master/Experimenter%20Companion/5GENESIS_Test_Cases_Companion_v1.0.pdf

	RSRP	TC_UBI_NBIoTRAN
MCPTT	time elapsed	TC_MCPTTAccessTime_MAL
	time elapsed	TC_MCPTTAccessTimeIncCallEstablishment_MAL
	time elapsed	TC_MCPTTMouthtoEarDelay
Application Specific KPIs		
Video Jitter	interarrival time variation	TC_JIT_VideoStreamJitter_MAL
IoT Application Latency	Packet Delay	TC_IoT_PacketDelayHTTPPOST_SUR
	Packet Delay	TC_IoT_PacketDelayMQTT_SUR_001
	Packet Delay	TC_IoT_PacketDelayCoAP_SUR
	Packet Delay	TC_IoT_PacketDelayMQTToverLORA_SUR
Video QoE	360° Live Video Streaming QoE	TC_360LiveVideoStreamingQoE_BER

3. MALAGA PLATFORM EXPERIMENTS

3.1. Overview

The goal of the second phase of experimentation in the Málaga Facility has been to validate two different infrastructure setups: the standard 5G NSA Option 3x deployment [12], and the experimental 5G setup based on the equipment provided by RunEL. Table 3-1 lists the KPIs evaluated in the second trial and summarizes the kind of evaluation measurements conducted.

The following tables present the available setups possible at the Málaga 5GENESIS Facility. Table 3-2 presents the 4G/LTE deployment configurations and Table 3-3 summarises the ones related to 5G. 5G setup numbers corresponds to the ones described in deliverable D4.5 [6].

Table 3-1 Primary 5G KPIs evaluated at the Málaga Platform in the second phase

KPI to be evaluated at the Málaga Platform according to DoA	Evaluated in Phase 2	Comment
Throughput	Yes	Based on iPerf
Latency	Yes	Based on RTT
Additional 5G KPIs evaluated at the Málaga Platform		
MCPTT Access time	Yes	-
MCPTT End-to-end Access	Yes	-
Content distribution streaming services: Video resolution, Time to load first media frame	Yes	-

Table 3-2 5GENESIS Málaga Platform deployed LTE setups detail

Deployment Parameters	LTE Products/Technologies Options		
ID	Setup 1.TRIANGLE	Setup 2.Indoor LTE	Setup 7. Indoor LTE VIM
Description	Legacy TRIANGLE testbed	Indoor E2E 4G setup	Indor E2E 4G setup in VIM
Core Cloud	No	No	Yes - OpenNebula
Edge Cloud	No	No	Yes - OpenNebula
# Edge Locations	1	1	1
Slice Manager	NA	Yes - Katana	Yes - Katana
MANO	NA	OSM v6	OSM v6
NMS	NA	TAP	TAP
Monitoring	NA	Prometheus	Prometheus
3GPP Technology	4G LTE+	4G LTE+	4G LTE+

3GPP Option	NA	NA	NA
Non-3GPP Technology	NA	NA	NA
Core Network	Polaris EPC	ATHONET Rel. 15 vEPC	Polaris EPC
RAN	OAI eNB	Nokia Flexizone picoBTS	Nokia Flexizone Small Cell
UE	COTS UE	COTS UE	COTS UE
Relevant Use Cases	TBD	Use Case 2	Use Case 3

Table 3-3 5GENESIS Málaga Platform deployed 5G setups detail

Deployment Parameters	5G Products/Technologies Options		
ID	Setup 3.Indoor 5G ECM	Setup 4.Indoor 5G REL	Setup 8. Full E2E 5G
Description	5G setup with ECM OAI solution	5G setup with RunEL solution	Indoor & outdoor E2E 5G (in progress)
Core Cloud	No	No	Yes - OpenStack
Edge Cloud	No	No	Yes - OpenNebula
# Edge Locations	NA	NA	1
Slice Manager	NA	NA	Yes - Katana
MANO	NA	NA	OSM v6
NMS	TAP	TAP	TAP
Monitoring	NA	NA	Prometheus
3GPP Technology	5G	5G	4G LTE+, 5G NSA
3GPP Option	NoS1	NoS1	NA
Non-3GPP Technology	NA	NA	NA
Core Network	No Core	No Core	ATHONET Rel. 15 vEPC (Setup 8.1) Polaris Rel. 15 EPC (Setup 8.2)
RAN	OAI eNB	RunEL eNB	Nokia Airscale System (indoor and outdoor)
UE	OAI UE	RunEL UE Emulator	COTS UE
Relevant Use Cases	TBD	TBD	Use Cases 1, 2, 3

The first setup is an NSA 5G NR deployment operated by UMA and located at the university campus. This deployment setup follows the NSA option 3x architecture [12] and supports two core options: Athonet EPC (Setup 8.1) and Polaris EPC (Setup 8.2), as shown in Table 3-3.

In the first setup we have executed standard test cases for measuring throughput and latency, in order to characterize the performance of the system after its deployment. In addition, MCPTT and content distribution streaming services test cases have been executed. These test cases are related on the use cases targeted in the Málaga Platform.

The second setup is experimental based on the RunEl RAN solution (i.e. 4.Indoor 5G REL) and for this testing session no integration with a 5G Core was available.

Due to the experimental nature of the second setup, custom test cases have been executed to measure latency.

3.2. Experiments and Results

3.2.1. Indoor & Outdoor E2E 5G Setup – Setup 8.1 Full E2E 5G

The system under test (SUT) includes 4 gNodes and 4 eNodes from Nokia and a 3GPP Rel.15 EPC. Two different Public Land Mobile Networks (PLMNs) are configured in the pilot, one is managed by Athonet and the other by Polaris. The data plane has been configured to use only the 5G data plane (data bearers are handled by gNB nodes). The commercial UE used during the testing has been Samsung Galaxy Note 10 (Exynos chipset). The UEs has been located in Line of Sight (LOS) and close proximity to achieve the maximum theoretical throughput of 286 Mbps for the discussed deployment. The most representative parameters of the 5G configuration applied are detailed in Table 3-4, which comprise the first stable scenario configured after the deployment of the network.

Table 3-5 provides the details of the 4G configuration applied in this setup. For comparison purposes, the tests have been also executed in 4G forcing in the UE the radio technology to LTE.

Table 3-4 5G NR Non-standalone mode network configuration

Band	n78
Mode	TDD
Bandwidth	40 MHz
Carrier components	1 Carrier
MIMO layers	2 layers
DL MIMO mode	2x2 Closed Loop Spatial Multiplexing
Modulation	256QAM
Beams	Single beam
LTE to NR frame shift	3 ms
Subcarrier spacing	30 kHz
Uplink/Downlink slot ratio	2/8

Table 3-5. 4G network configuration

Band	B7
Mode	FDD
Bandwidth	20 MHz
Carrier components	1 Carrier
layers	4 layers
DL MIMO mode	4x4 Closed Loop Spatial Multiplexing
Modulation	256QAM

3.2.1.1. Throughput

This test is devoted to the measurement of the throughput in the downlink between the main compute node and a 5G UE. The test has been executed automatically via the 5GENESIS Coordination Layer, iPerf TAP plugins and the iPerf agents developed in WP3.

The traffic originates at the main compute node connected to the core network (CN) and is received at the 5G UE. There is a direct line of view between the 5G UE and the gNodeB. The throughput obtained is close to the theoretical maximum (286 Mbps) for the deployment setups described in Table 3-3. The results of the experiment are depicted in Figure 3-1. In light of the results, we can conclude that the performance of the scenario and setup under test has been validated in terms of throughput. The details of the test case executed, and the statistical results are included in Annex Table A-1.

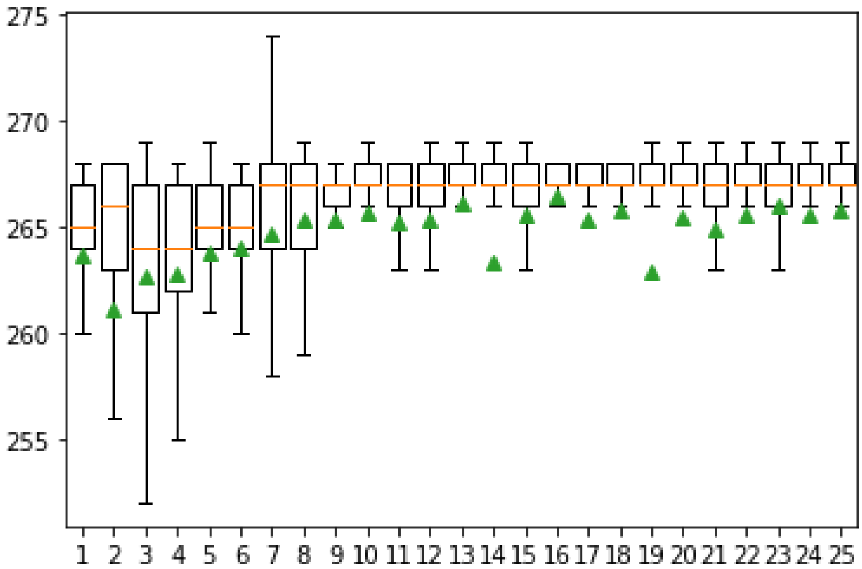


Figure 3-1 5G NSA 3x MIMO 2x2 TDD 40 MHz 256 QAM throughput

Figure 3-2 shows the results of executing the same test in the LTE deployment described in Table 3-5. The throughput is lower than in the 5G scenario, however the difference is not high due to the different number of MIMO arrays used in those deployments, i.e. the 5G scenario has a MIMO 2x2 RAN configuration whilst the 4G scenario is configured as MIMO 4x4. In this sense, the two setups are not identical for direct comparison.

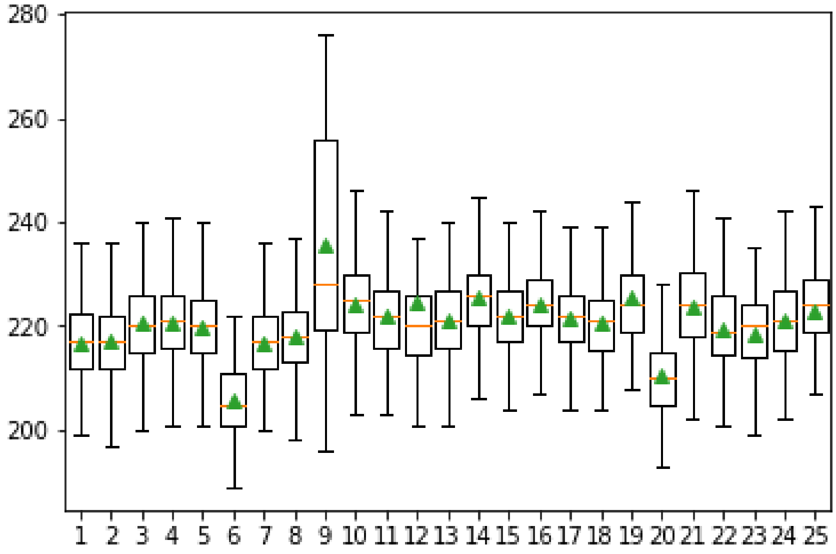


Figure 3-2 LTE 20 MHz MIMO 4x4 256 QAM Throughput

3.2.1.2. Round trip time

This test is devoted to the measurement of the RTT between a 5G UE and the Packet Data Gateway of the EPC. The test has been executed automatically via the 5GENESIS Coordination Layer, ping TAP plugin and the ping agent developed in WP3.

The ping messages are initiated by the UE. There is a LOS between the 5G UE and the gNodeB. The results of the experiment in the 5G scenario described in Table 3-4 are depicted in Figure 3-3. The most representative parameters of the 5G configuration applied are detailed in Table 3-4, this is the first stable scenario configured after the deployment of the network. The mean RTT obtained for the network configuration is around 12s.

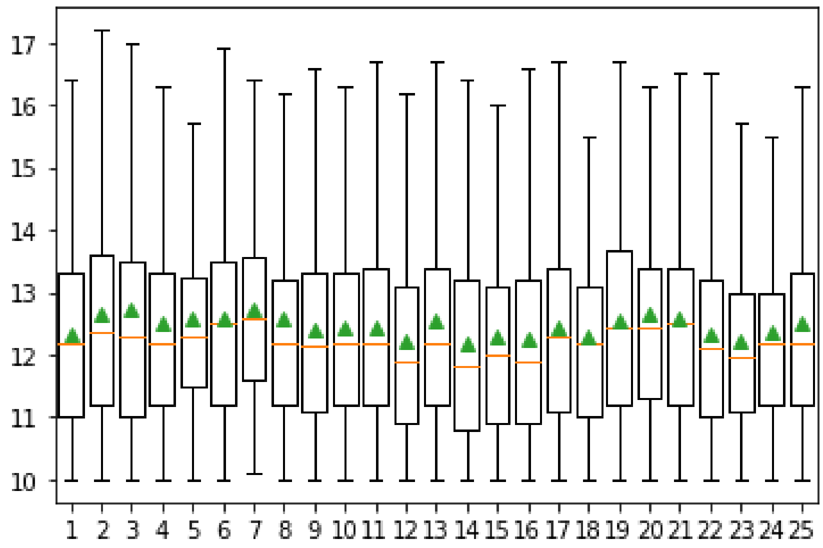


Figure 3-3 5G NSA 3x MIMO 2x2 TDD 40 MHz 256 QAM RTT

The achieved value is lower than in the 4G setup, as shown in Figure 3-4 which is close to 33 ms for the setup described in Table 3-5. The detailed results are presented in Annex Table A-2.

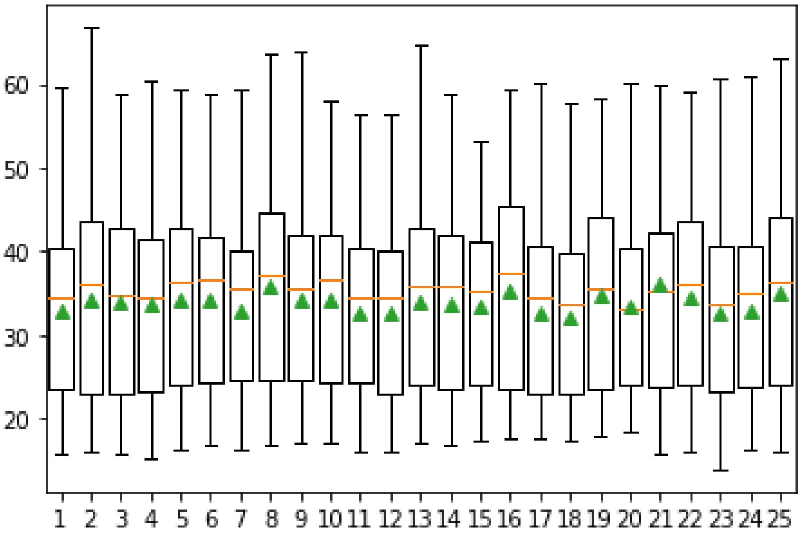


Figure 3-4 LTE 20 MHz MIMO 4x4 256 QAM RTT

3.2.1.3. Content distribution streaming services

The test case executed in this subsection has been specified in “D2.6 Final Test Scenario and Test Specifications” from the TRIANGLE project [13]. As one of the use cases targeted in the Málaga Platform is video surveillance, we have used this test case to evaluate the performance of content distribution streaming services over 5G.

Figure 3-5 shows the time to load first media frame during video streaming sessions. The delay is a key KPI impacting in the QoE and a critical KPI in public safety applications. Figure 3-5 shows the improvement obtained in 5G.

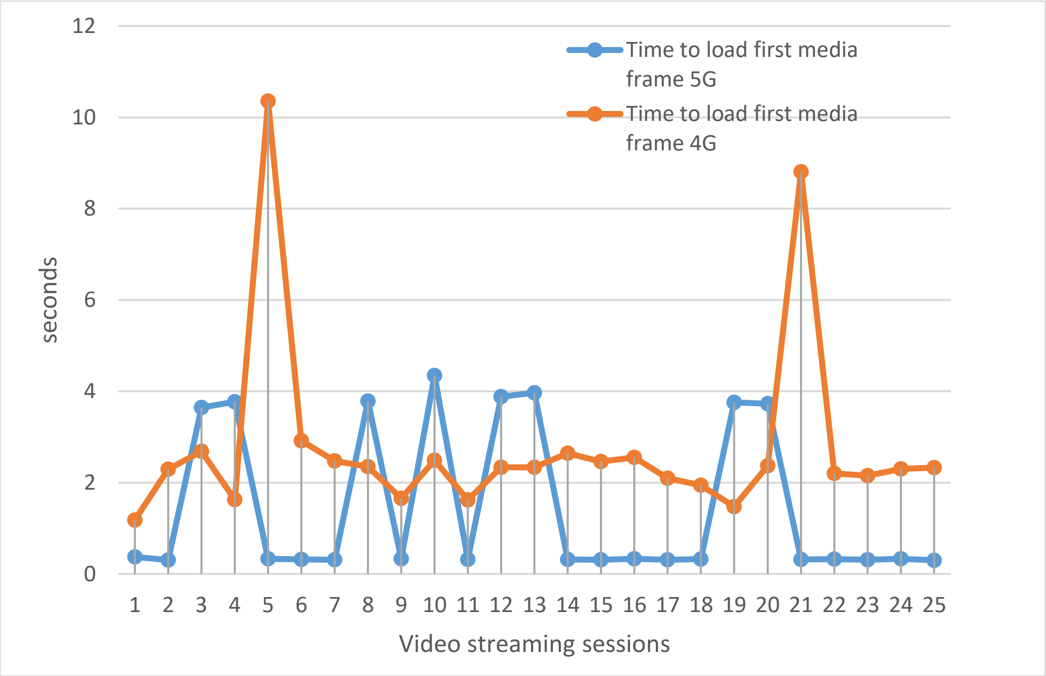


Figure 3-5 5G vs 4G Time lo load first media frame

Video resolution is also a key KPI for video streaming service. Figure 3-6 shows video resolutions obtained during 25 DASH streaming sessions with a duration of 3 minutes each one of them.

The results obtained also demonstrated a clear improvement of the resolution in 5G. The detailed results are included in Annex Table A-3.

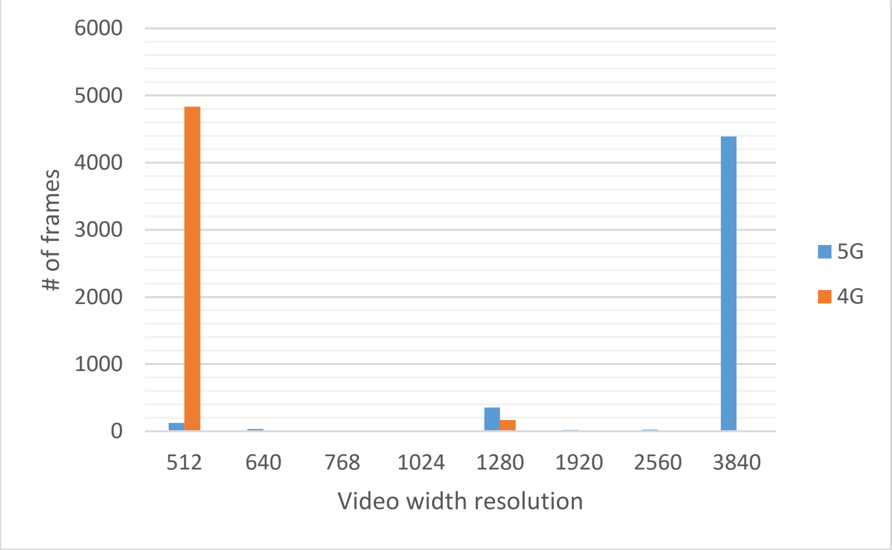


Figure 3-6 5G vs 4G scenario Video resolution

3.2.1.4. MCPTT

Both 5G setups 8.1 and 8.2 (repeating the experiments with both Athonet and Polaris Release 15 NSA EPCs) have been used to evaluate the MCPTT KPI. The configuration used has been the same than that described previously in Section 3.2.1, including the radio parameters depicted in Table 3-4. Results are present in Figure 3-7.

Many different experiments have been performed to measure this KPI. The MCPTT KPI has been evaluated for two different MCS services: Nemergent MCS service, and Airbus MCS service. For the Nemergent service, the MCPTT KPI has been subdivided into MCPTT Access Time and MCPTT E2E Access Time while, for the Airbus service, just MCPTT Access Time has been evaluated (due to the Airbus MCS service app not providing enough information to process MCPTT E2E Access Time). All those experiments have been performed for both 4G and 5G data connection (forcing the UEs to use 4G or 5G from their own network settings, but with the same setup 8 for 4G and 5G).

This results in a total of 12 experiments, allowing the evaluation of this KPI with confidence for both MCS services in 4G and 5G. As a summary, the multiple experiments executed for this KPI attend to the use of:

- 5G and 4G for data plane (forcing it at the UEs, no change in setup 8).
- Nemergent MCS service and Airbus MCS service.
- Athonet Rel. 15 NSA EPC and Polaris NetTest Rel. 15 NSA EPC.
- MCS Access Time and MCS End-to-end Access Time for Nemergent MCS Service, just MCS Access Time for Airbus MCS service.

Nemergent MCS service

In this second cycle, regarding MCPTT validations we take into account Malaga Facility evolution related to availability of commercial 5G NSA UEs, commercial NOKIA eNB/gNB, improvements in the 4G/5G cores and improvements in the deployed MCPTT service itself among others.

In cycle one, the MCPTT results were very satisfactory and provided measurements for Access Time with values near to 50 ms, while for E2EAccess Time the values hovered 250 ms approximately. The difference among the measurements was coherent, since E2E included the time for MCPTT call establishment and then the token granting time, while Access Time just measured the time for the token to be granted (as defined in [14] for Access Time and End-to-End Access Time).

Current results demonstrate the importance of the platform and service evolution showing even lower values in a consistent way. The second cycle depicts an average Access Time of 28.82ms and 27.95ms for the 4G core of Athonet and Polaris respectively, while involving 5G cores the values go down to 17.68 ms and 16.72 ms. For E2E Access Time, the results are 137.94 ms and 145.87 ms for Athonet and Polaris with 4G cores and 138.15 ms and 128.24 ms. The detailed results are included in Annex Table A-4, Annex Table A-5, Annex Table A-6 and Annex Table A-7 for the Access Time and in Annex Table A-8, Annex Table A-9, Annex Table A-10 and Annex Table A-11 for the end-to-end Access Time.

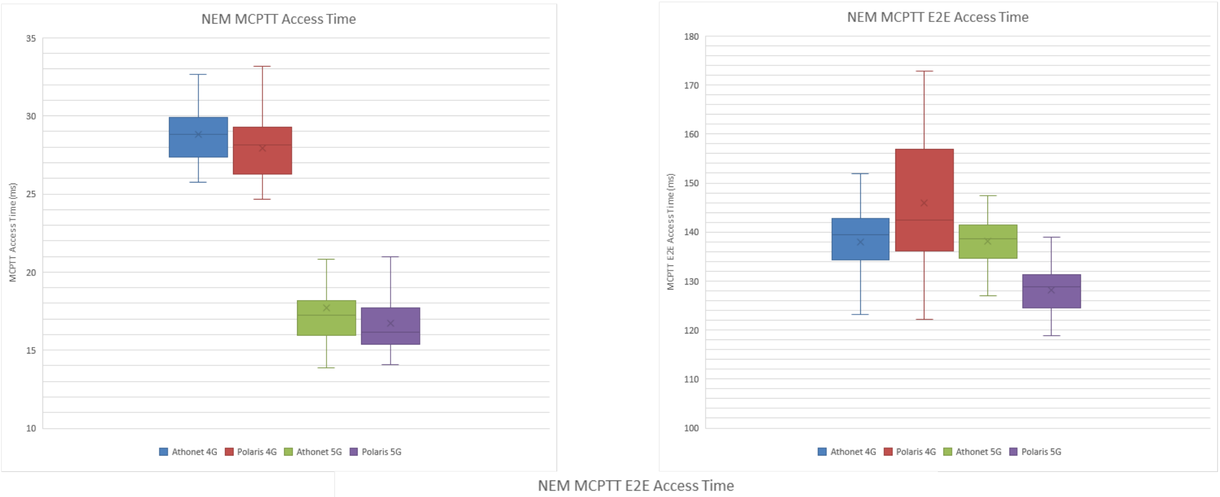


Figure 3-7 a) MCPTT Access Time for NEM MCS; b) MCPTT End-to-end Access Time for NEM MCS

Considering the standardized thresholds for each KPI, the good values obtained well below the max set threshold (300 ms for Access Time and 1000 ms for E2E Access Time in 3GPP TS 22.179 [14]) and the difference from the previous cycle, the results clearly show four important facts:

1) The testing environment (i.e. setup 8) does not introduce any additional delays, therefore it provides the perfect ground to perform reliable tests on technology and services.
2) The tested MCS service is efficient in a way that the service itself only consumes less than a third part of the total threshold to achieve the measured task. The cycle 2 values also manifest a greater gap between [11] the obtained KPIs with non-loaded network and the official standardized thresholds, giving more room for hosting a greater number of active and parallel mission critical subscribers while ensuring service Quality of Service (QoS).

3) The involved 5G equipment greatly improves the results, showing a clear platform evolution and demonstrating once again its suitability to host services, either for prototyping, benchmarking or adaptation to 5G procedures.

Also, the E2E Access Time exhibits a similarity while using 4G or 5G with the NSA cores of the experiments, which identifies the 5G core (NSA and SA) as a very clear candidate that needs evolution for last cycle.

Results represented previously in show a clear impact of the Málaga Platform evolution, especially regarding its 5G NSA setup, which will still be improved for the next cycle in order to add further enhancements and capabilities including a full 5G SA setup.

Airbus Agnet MCS Service

Results for the Airbus MCS MCPTT experiments are very similar to the previously presented for Nemergent MCS. In the case of the Airbus Access Time experiments, they show average times of 40.65 ms and 35.96 ms using 4G with the cores of Athonet and Polaris respectively, and for 5G 29.01 ms and 28.10 ms. The acquired results are few milliseconds higher than those in the Nemergent experiments, probably due to the Airbus MCS service processing time being higher than in Nemergent's case. The detailed results are available in Annex Table A-8, Annex Table A-9, Annex Table A-10 and Annex Table A-11.

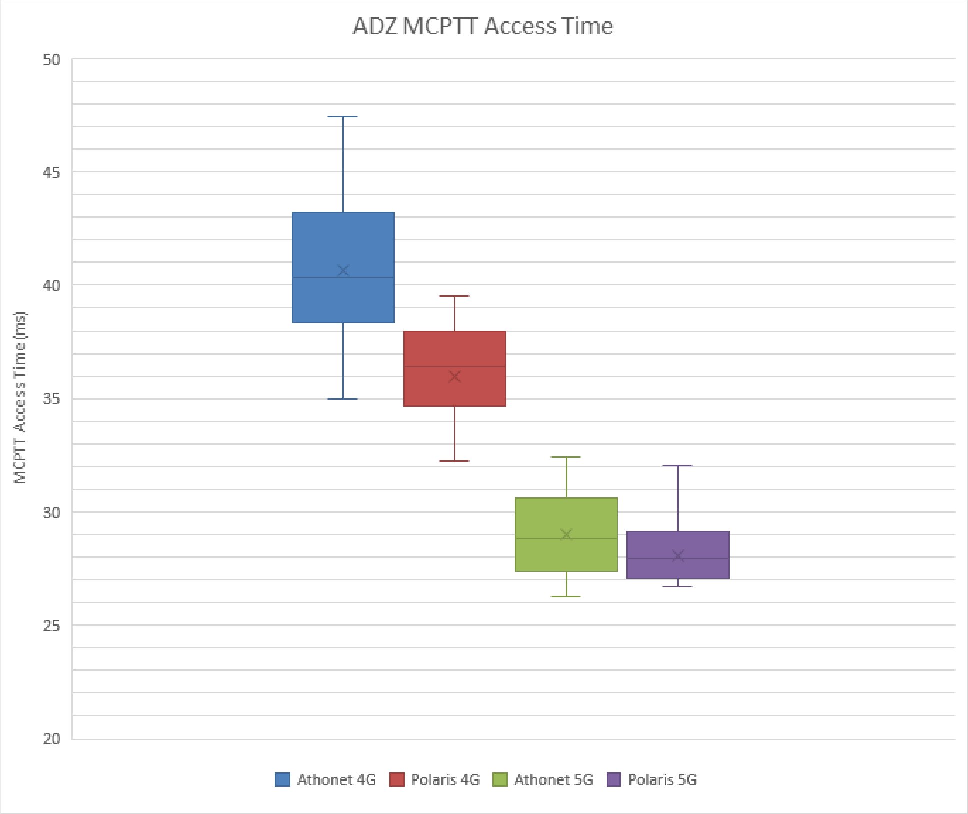


Figure 3-8. MCPTT Access Time for ADZ MCS

The results obtained are still very good in comparison with the standardized threshold for the Access Time KPI of 300 ms (1000 ms for End-to-end Access Time) defined in [14]. This supports and strengthen the conclusions previously mentioned for the Nemergent MCPTT experiments.

3.2.2. RunEL 5G RAN setup – Setup 4. Indoor 5G REL

3.2.2.1. Latency

The RunEL 5G setup corresponds to setup 4 described in deliverable D4.5 [6]. A block diagram representing the exact setup used for this experiment can be seen in Figure 3-9. This setup does not include an EPC, but only the 5G radio prototype from RunEL. The results are summarized in Annex 9.A.1.2.

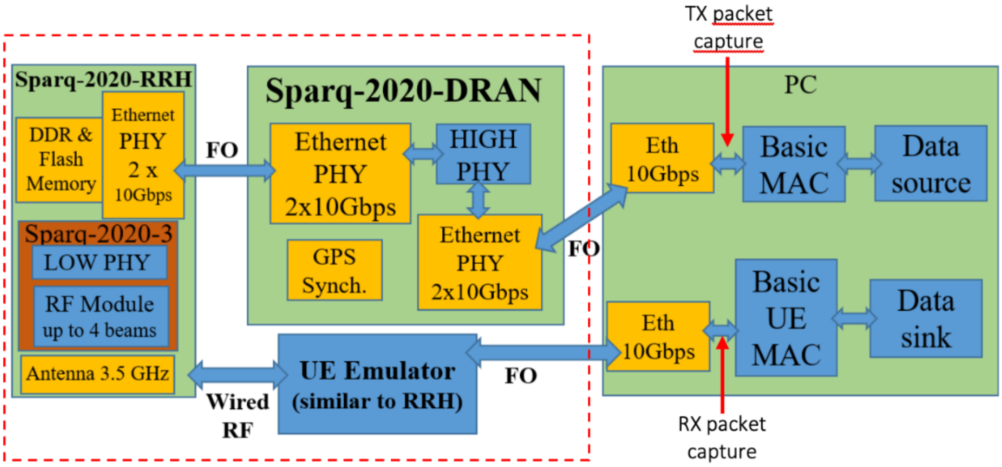


Figure 3-9 RunEL setup at Málaga Platform for PHY delay measurements

This experiment allowed to measure the downlink latency present in the air or radio interface, which includes just the PHY layer of the radio stack. For that purpose, we must measure the latency limited to the part of the setup highlighted in Figure 3-9.

Considering that in this setup the MAC layer is an independent software running in a PC for both data source and destination, it is possible to isolate the PHY layer of the setup. Firstly, we captured traffic as seen in Figure 3-9, and then the same was done to remove the delay introduced by the PC and its network interfaces as seen in Figure 3-10. This way we could precisely calculate the PHY layer latency.

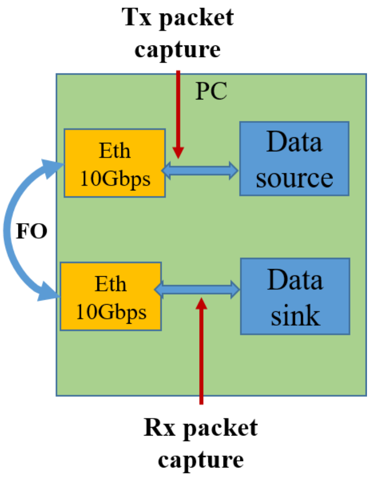


Figure 3-10. Setup and measurement points without SUT for PHY latency testcase

Regarding the results (detailed in Annex Table A-16), the mean value observed is 1.408 ms, which is in line with the target value of 2 ms for latency on the air interface, as shown in Table 3-1 of Deliverable D6.1 [11], which summarises 5G-PPP KPIs and target values. The results also demonstrate that PHY Layer latency KPI is very stable with the RunEL setup, varying just some microseconds among different iterations.

4. ATHENS PLATFORM EXPERIMENTS

4.1. Overview

During 5GENESIS Trials and Experimentation Cycle 2, the Athens Platform focused on conducting experiments on its commercial 5G NSA systems, based on Amarisoft Classic Callbox (RAN and CN) and Athonet EPC Rel.15 CN. The setups correspond to 5G.4.Option3 (Amarisoft RAN and Amarisoft CN) and 5G.5.Option3 (Amarisoft RAN and Athonet CN). In addition, we used the 5GENESIS Coordination Layer, which released as open-source under the name “Open 5GENESIS Suite”, to perform the experiments (Portal-OpenTAP) and the 5GENESIS Analytics Framework to perform Statistical Analysis of the recorded data.

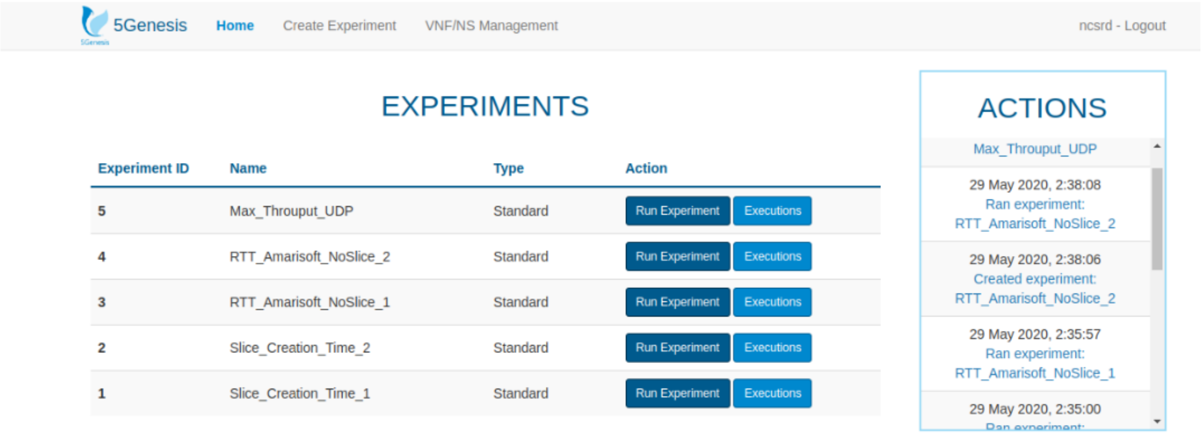
We also used UMA’s iPerf, Ping and Resource Agents for recording data on the 5G COTS UEs, while in some experiments we utilized instead the Android Application “MNL Metrics Tool” (refer to Figure 4.10 for further details), developed by the Media Networks Laboratory of the National Centre of Scientific Research "Demokritos" (NCSRD), providing radio metrics recording, Ping and iPerf utilities. All applications send the recorded data in InfluxDB, in order to perform statistical analysis.

Table 4-1 lists the KPIs evaluated in the second trial and summarizes the kind of evaluation measurements conducted.

Table 4-1 KPIs evaluated in the Athens Platform during Phase 2

KPI to be evaluated at the Athens Platform according to DoA	Evaluated in Phase 2	Comment
Ubiquity	No	Not scheduled for Phase 2
Latency	Yes	
Capacity	Yes	Phase 2 focused on Throughput measurements (see below)
Service creation time	Yes	-
Additional 5G KPIs evaluated at the Athens Platform	Evaluated in Phase 2	Comment
RTT	Yes	-
Throughput	Yes	-

All experiments were conducted using the 5GENESIS Experimentation Methodology. The experiments include Throughput, E2E RTT, Latency (one-way delay) and Service Creation Time. We also provide variations of these experiments, by conducting measurements in various cell locations (mid-edge, cell-edge) and under concurrent network traffic in the E2E RTT Test Case, thus providing more insight on the behaviour of real 5G NSA networks.



EXPERIMENTS

Experiment ID	Name	Type	Action
5	Max_Throuput_UDP	Standard	Run Experiment Executions
4	RTT_Amarisoft_NoSlice_2	Standard	Run Experiment Executions
3	RTT_Amarisoft_NoSlice_1	Standard	Run Experiment Executions
2	Slice_Creation_Time_2	Standard	Run Experiment Executions
1	Slice_Creation_Time_1	Standard	Run Experiment Executions

ACTIONS

Max_Throuput_UDP

29 May 2020, 2:38:08
Ran experiment:
RTT_Amarisoft_NoSlice_2

29 May 2020, 2:38:06
Created experiment:
RTT_Amarisoft_NoSlice_2

29 May 2020, 2:35:57
Ran experiment:
RTT_Amarisoft_NoSlice_1

29 May 2020, 2:35:00

Figure 4-1 NCRSD Portal interface

In the following sections, we report the KPIs per setup in the following order:

- Amarisoft RAN – Amarisoft CN:
    - Throughput.
    - E2E RTT with different packet sizes.
    - E2E RTT with background traffic.
    - E2E RTT in different cell locations.
    - Latency (one-way delay).
    - Service Creation time.
- Amarisoft RAN – Athonet CN:
    - Throughput.
    - E2E RTT with different packet sizes.

Table 4-2 Athens Platform 5G Deployment Configurations

Deployment Parameters	Deployment Flavors			
ID	5G.1.noS1	5G.2.noS1	5G.3.Option3	5G.4.Option3
Status	Under deployment	Planning	Operational	Operational
Description	No Core and NR proprietary	No Core, Vendor NR	Vendor Core/gNB	Vendor All-in-one deployment
Core Cloud	NA	NA	NA	NA
Edge Cloud	NA	NA	NA	NA
# Edge Locations	1	1	1	1
WAN/Network	NA	NA	SDN	SDN

Slice Manager	NA	NA	NA	NA
MANO	NA	NA	NA	NA
NMS	NA	NA	NA	NA
Monitoring	NA	NA	Prometheus	Prometheus
3GPP Technology	5G	5G	5G	5G
3GPP Option	noS1	noS1	NSA	NSA
Non-3GPP Technology	NA	NA	NA	NA
Core Network	NA	NA	Athonet EPC	Amarisoft EPC/5G Core
RAN	OAI gNB	RunEL DRAN	Amarisoft gNB (SDR)	Amarisoft gNB (SDR)
UE	OAI nr-UE (SDR)	OAI nr-UE (SDR)	Samsung A90 5G	Samsung A90 5G

4.2. Experiments and Results

4.2.1. Amarisoft RAN – Amarisoft CN (5G.4.Option3)

Figure 4-2 illustrates the test setup for the experiments we conducted in the Amarisoft RAN-Amarisoft CN 5G NSA setup in Athens Platform. In all experiments, traffic flows between Endpoints 1 & 2, namely a Samsung A90 5G (SM-A9080) and a commodity Dell Laptop.

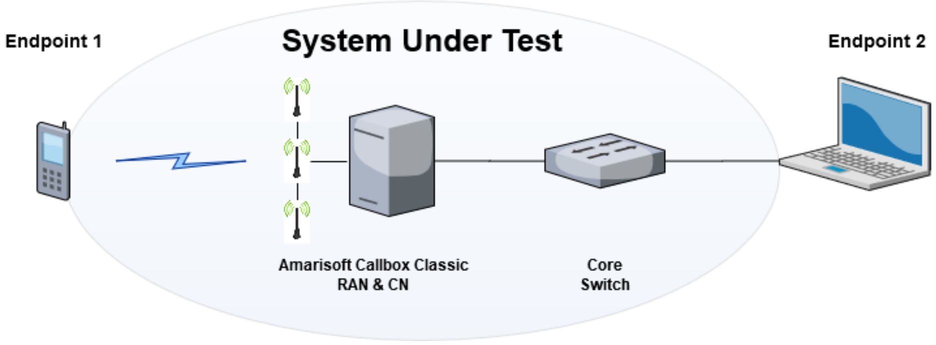


Figure 4-2 Amarisoft RAN – Amarisoft CN testbed setup

4.2.1.1. Throughput

During the experiment, the radio conditions were excellent and stable, as recorded in the UE by UMA's Resource Agent and are illustrated in Figure 4-3. The radio metrics captured were RSRP, RSRQ and RSSI with the following average values: RSSI= -51.00 +/- 0.00 dBm, RSRP= -70.48 +/- 0.20 dBm, RSRQ=-6.51 +/- 0.21 dB.

The radio configuration of the 5G cell in this setup corresponded to 50 MHz bandwidth, 2x2 MIMO, TDD, Band n78, 256 QAM DL, resulting in a theoretical throughput of 477 Mbps. Before running the experiment, it was important to adjust the UDP data rate, in order to achieve an acceptable level of packet loss. The selected UDP data rate for the experiment was eventually adjusted to 377 Mbps, providing a mean packet loss approximately around 1% during the experiment, as reported by the iperf2 probes.

The detailed results of the primary metric, its first order statistics and the complementary metrics of the experiment are presented in Annex Table A-17. The average throughput was calculated to 369.27 +/- 0.61 Mbps, corresponding to a decrease of 2,2% of the selected UDP data rate. However, it is important to note that the 95th percentile reports a value of 373.06 +/- 0.13 Mbps, showing that 95% of the recorded values were below 373.06 Mbps. Percentiles are more effective in describing the performance of a system contrary to the average, as they can capture the real distribution of the data.

Another important parameter worth noting is the minimum throughput value of 285.48 +/- 11.69 Mbps. By inspecting the recorded metrics, we noticed that these minimum values were reported at the beginning of some iterations out of the total 25, along with the highest packet loss values. This behaviour may be explained by the buffers of the system that filled up as a result of the previous iterations. It is also worth noting that the 5th percentile corresponds to 369.08 +/- 2.99 Mbps, clearly showing that 5% of the throughput values were below 369.08 Mbps and that the minimum values recorded could be considered outliers. These values were also most probably the reason that the upper bound of the 95th percentile of packet loss was calculated to 1.75%.

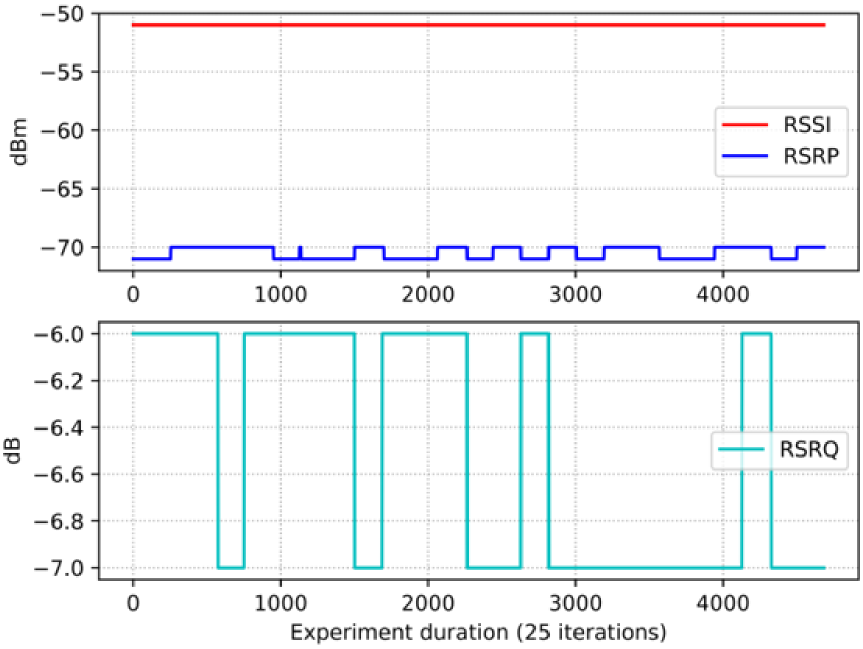


Figure 4-3 Typical radio conditions during throughput experiments

4.2.1.2. Round-Trip Time

This test case evaluates the impact of the packet size on the E2E RTT metric. Different packet sizes refer to different applications, so we are able to evaluate the system's response on various use cases, ranging from file sharing to audio and video traffic streaming. The packet sizes used are 32, 64, 128, 512 and 1500 bytes.

During the experiment, the radio conditions were excellent and stable, as recorded in the UE by UMA's Resource Agent, and are illustrated in Figure 4-4. The radio metrics captured were RSRP, RSRQ and RSSI and their detailed results are reported in Annex Table A-18. As an indication, the typical radio conditions correspond to RSSI= -51.00 +/- 0.00 dBm, RSRP= -71.80 +/- 0.22 dBm, RSRQ= -6.67 +/- 0.16 dB.

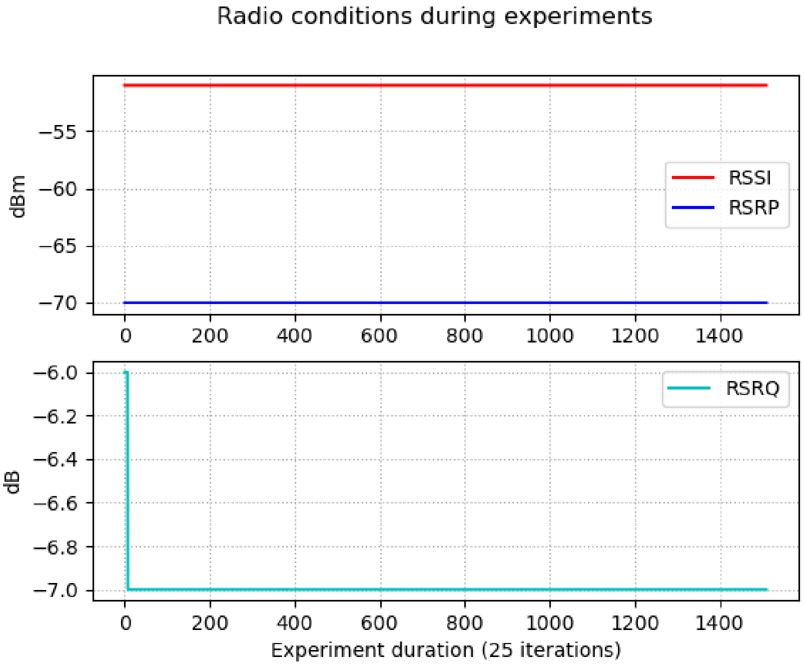


Figure 4-4 Typical radio conditions during E2E RTT experiments

The radio configuration of the 5G cell in this setup is 50 MHz Bandwidth, 2x2 MIMO, TDD, Band n78, 256 QAM DL, while the 4G cell provided 10 MHz Bandwidth, 2x2 MIMO, FDD in Band 1. The reported average E2E RTT of a single connected 5G COTS UE is 34.66 +/- 0.24 ms with 64 bytes packet size on an empty channel without background traffic (see Figure 4-5). The 95th percentile E2E RTT is 47.99 +/- 0.49 ms. It is also important to note that all ping requests were successful, resulting in an average ping success ratio of 1.00, due to low network load and stable radio conditions throughout the experiment.

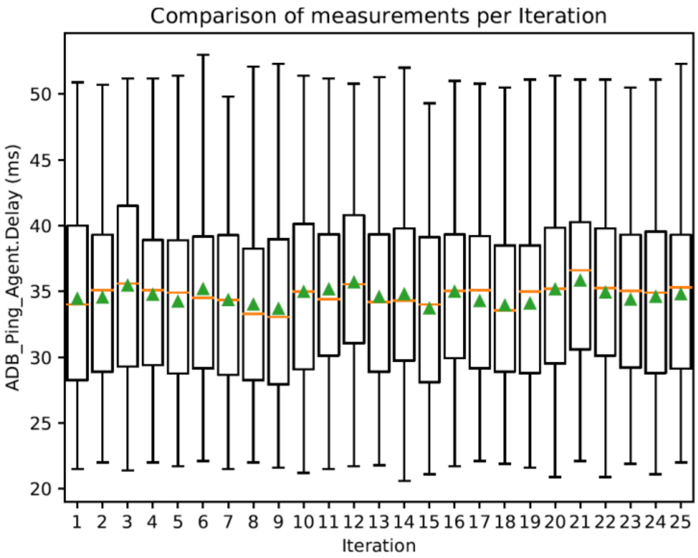


Figure 4-5 E2E RTT per iteration (64 bytes packet size), provided by SRL's Analytics Framework

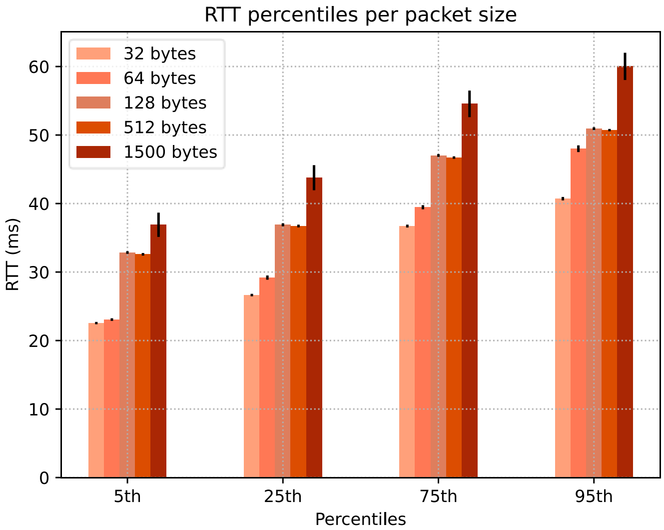


Figure 4-6 E2E RTT percentiles per packet size

As expected, the RTT in 5G NSA networks is comparable to that of the the 4G case, since the 5G NR Cell is anchored to an existing 4G deployment and uses the same CN (EN-DC). This clearly shows that initial 5G NSA deployments are suitable for supporting eMBB applications requiring higher throughput than the one 4G networks provide. 5G NSA deployments allow for quick rollouts and cost-effective coverage, leveraging the existing 4G infrastructure. However, the very low latency required by many use cases will only be achieved in 5G SA deployments, which are designed to support the uRLLC case.

The detailed results of each packet size are provided in Annex Table A-18. It is clear that packet size affects the E2E RTT, providing a range of average values from 31.68+/-0.16 ms (32 bytes) to 48.98 +/- 1.81 ms (1500 bytes). The bar chart in Figure 4-7 provides an overall overview of the percentiles of E2E RTT per packet size. It is important to note that E2E RTT for 128- and 512-bytes packet sizes almost overlap, indicating that the buffers of the SUT are optimized for traffic of such level. In addition, the standard deviation of E2E RTT is comparable for all packet sizes, ranging from 5.97 +/- 0.10 ms to 7.28 +/- 0.15 ms, indicating the stability of network conditions in the SUT.

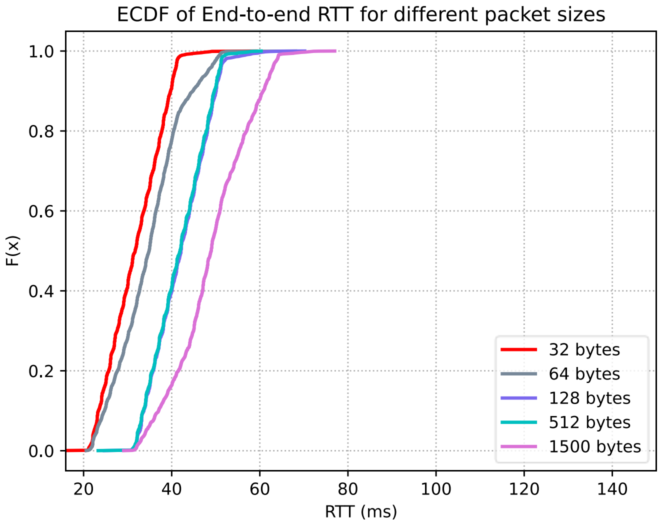


Figure 4-7 ECDF of E2E RTT for different packet sizes in 5G.4.Option3 setup

We also provide the Empirical Cumulative Distribution Function (ECDF) of the data gathered in these experiments in Figure 4-7. The ECDF clearly presents the performance degradation as the packet size increases. We also notice here the E2E RTT overlap between 128 bytes and 512 bytes in our SUT.

4.2.1.3. RTT with background traffic

In addition, we conducted further experiments on E2E RTT using background traffic. In this case, we measured E2E RTT while transmitting UDP traffic via iperf2 with 377 Mbps data rate. As expected, there was an overall increase in the E2E RTT metric compared to the previous experiment without background traffic. Specifically, the average E2E RTT was 37.84 +/- 1.21 ms, the median 37.47 +/-1.19 ms and the 95th percentile 53.21 +/- 2.55 ms. In this case, there was 9.17%, 7.89% and 10.88% increase in each metric respectively, compared to not transmitting any background traffic. All first order statistics are presented in detail in the test report in Annex Table A-19.

It is important to consider the level of background traffic transmitted throughout this experiment, corresponding to almost 100% utilization with packet loss approximately around 1%. However, there were not any ping failures (Ping Failed Ratio was 0.00), while the increase in E2E RTT can be described as moderate. This is due to having only one UE connected to the network, utilizing all available resources under stable and excellent radio conditions, as shown in Figure 4-8.

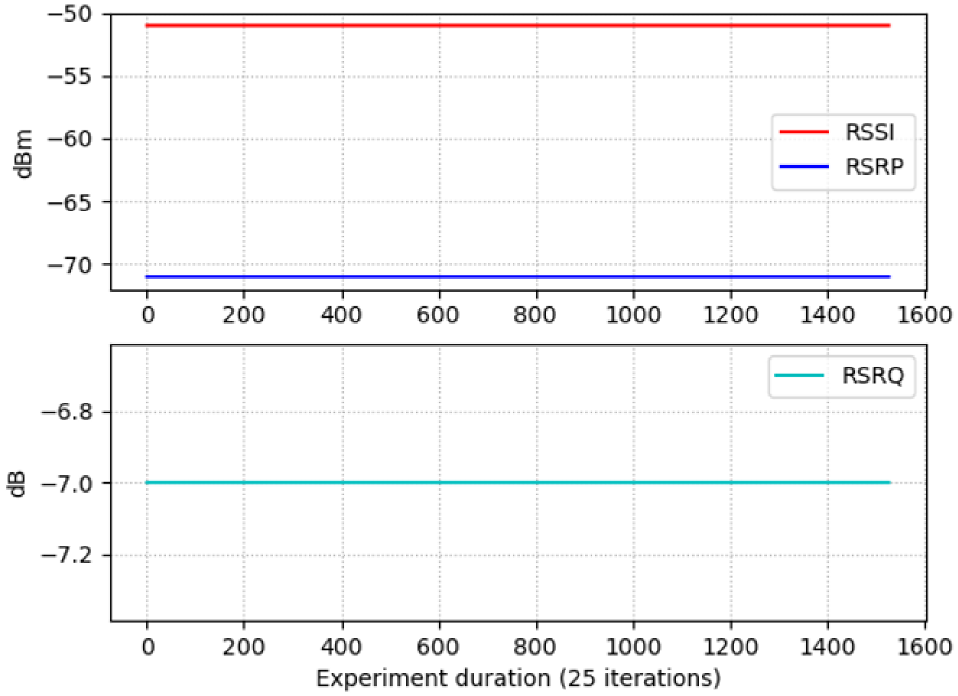


Figure 4-8 Radio conditions during E2E RTT experiment with background traffic

4.2.1.4. E2E RTT in relation to Radio Link Quality

This test case evaluates the E2E RTT in various cell locations, where the radio link quality ranges from excellent to edge conditions.

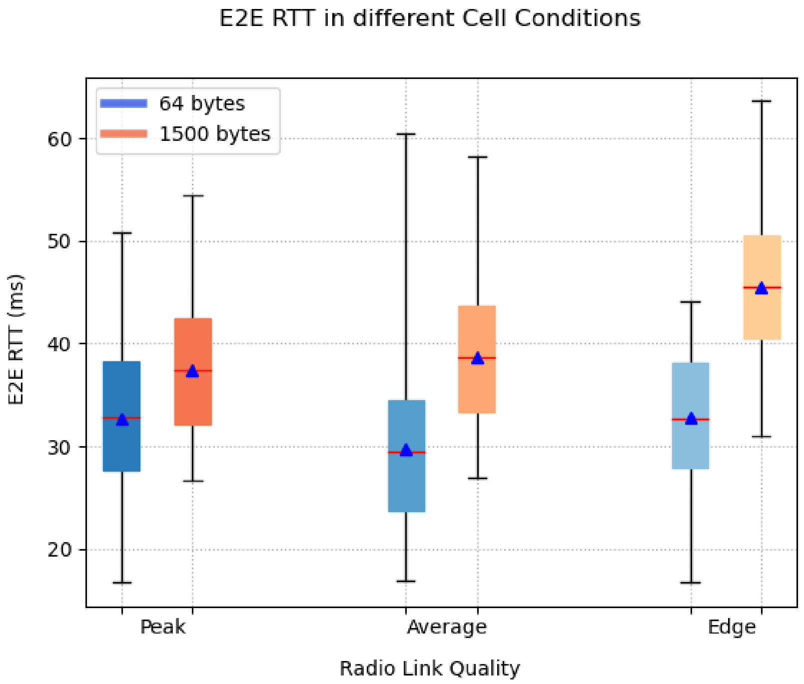


Figure 4-9 E2E RTT in different cell locations

Figure 4-9 illustrates the mean E2E RTT for 64- and 1500-bytes packet size in three different cell locations. The radio link quality has low impact on E2E RTT for low packet sizes, in agreement with the reported results in the relevant experiment published in the Deliverable "5G Pre-Commercial Networks Trials Major Conclusions" by NGMN Alliance[2]. In addition, there is a moderate increase of 21.76% in the E2E RTT on 1500 bytes as we move towards the cell edge. As a result, the radio link quality has moderate impact on large packet sizes. The detailed results are provided in Annex Table A-20.

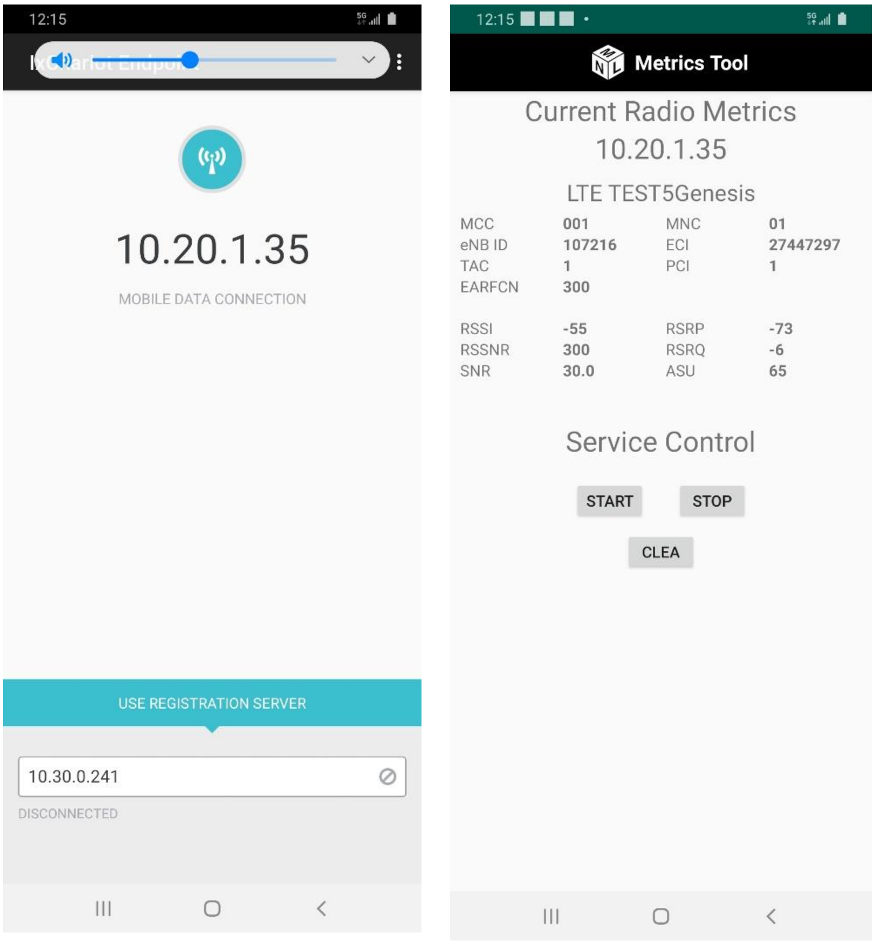
4.2.1.5. Latency (one-way delay)

In 5GENESIS, Latency (one-way delay) is defined as the time between the transmission and the reception of a data packet at application level. According to *TC_LAT_e2eAppLayerLatency*, the measurement methodology argues that the traffic profile is application-based, so we employed Real-time Protocol (RTP) (10 Mbps) traffic to measure one-way delay in the network.

In this experiment, we used IXIA's IxChariot Traffic Generator, which provides a handful of traffic profiles and generates application-specific statistics, such as one-way delay, Jitter and Throughput. IXIA provides software probes that are installed on the measurement endpoints and allow registration in IxChariot's Registration Server. All nodes of the network are synchronized to the same NTP server (Stratum 1) of the laboratory, to ensure accuracy between their clocks. We also used MNL's Android Application "MNL Metrics Tool" (Figure 4-10) to record radio metrics on the COTS 5G UE and store them on InfluxDB for further processing.

---

[2] https://www.ngmn.org/wp-content/uploads/Publications/2020/20200130_NGMN-PrecomNW_Trials_Major_Conclusions.pdfhttps://www.ngmn.org/wp-content/uploads/Publications/2020/20200130_NGMN-PrecomNW_Trials_Major_Conclusions.pdf



12:15

10.20.1.35

MOBILE DATA CONNECTION

USE REGISTRATION SERVER

10.30.0.241

DISCONNECTED

12:15

Metrics Tool

Current Radio Metrics

10.20.1.35

LTE TEST5Genesis

MCC	001	MNC	01
eNB ID	107216	ECI	27447297
TAC	1	PCI	1
EARFCN	300		
RSSI	-55	RSRP	-73
RSSNR	300	RSRQ	-6
SNR	30.0	ASU	65

Service Control

Figure 4-10 Ixia's IxChariot Endpoint and MNL's Metrics Tool

The radio metrics captured are the RSSI, RSRP and RSRQ. Their values indicate excellent and stable conditions throughout the experiment and are illustrated in Figure 4-11:

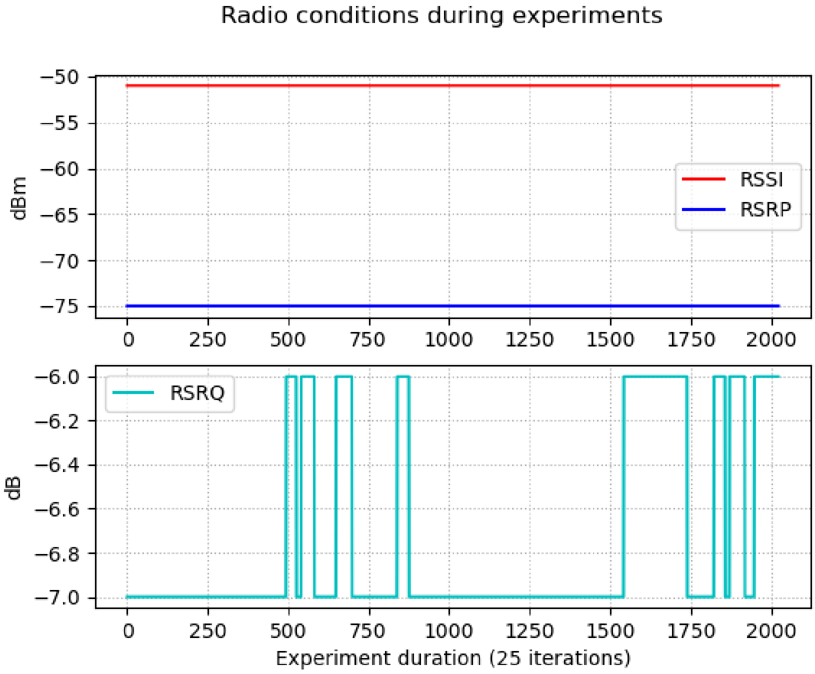


Figure 4-11 Radio Conditions during Latency experiments

The misconception that one-way delay is half of the RTT is really common. This claim is not always correct, since latency may vary between the DL and the UL, due to network topology (symmetric/asymmetric network) or queuing delays. For this reason, we conducted separate measurements for the DL and UL Latency, in order to show potential differences. It is important to note though that our SUT is a symmetric network (packets follow the same route on the DL and UL), while there is no severe network load.

As shown in the test results below, the average DL one-way delay is 21.37 +/- 0.28 ms and the UL one-way delay is 16.15 +/- 0.15 ms and as explained above, there is moderate difference between these two values. Another note is that their sum is comparable to the E2E RTT metric we measured in the previous experiments (64 bytes packet size, Mean E2E RTT = 34.66 +/- 0.24 ms). The endpoints use IXIA's proprietary clock synchronization algorithm for estimating the clock difference between the endpoints and we report the estimated error in each measurement (DL: 5.68 +/- 0.24 ms, UL: 5.78 +/- 0.18 ms) to provide reference on the accuracy of the clock difference. It is important to note that one-way delay is reported for the specific type of traffic (RTP 10 Mbps) and should not be generalized to different traffic profiles.

Although not required by the specific test case, we are also reporting the Jitter, Throughput and Delay Factor as complementary metrics for the sake of completeness. Specifically, according to IxChariot's documentation, the Delay Factor evaluates the size of the jitter buffer which would be required to eliminate the video interruptions due to network jitter and its measured values (DL: 29.14+/- 0.17 ms, UL: 38.28 +/- 0.13 ms) are in the acceptable levels of 9-50 ms[3]. The detailed results are available in Annex Table A-21.

4.2.1.6. Service Creation Time

Service instantiation and network slicing inside shared compute virtual infrastructures is a key target of the Open5GENESIS experimentation Framework [15]. In this respect, the Service Creation Time experiments are essential to evaluate the platform's performance and capabilities.

In Phase 2, the 5G E2E connectivity test case is used for experimentation inside the Athens Platform. The test case aims to provide network connectivity to a specific location using a 5G Mobile Network deployed and configured inside a sliced infrastructure. Initial request to start the experiments is sent by the 5GENESIS Portal to the Experiment Life Cycle Manager (ELCM) that is responsible for applying the 5GENESIS Experimentation Methodology of 25 consecutive service instantiations followed by results collection. Responsible for the slice and service creation is the Slice Manager along with the Network Management System (NMS). Additional components used in Athens Platform include the Virtual Infrastructure Manager (VIM) and the Network Function Virtualization Orchestrator (NFVO), realised by OpenStack and OSM respectively. More information about the Slice Creation process and the Network Slice Template (NST) can be found in deliverable D3.3 [16].

The experiments provide metrics regarding the duration of the process, beginning from the moment a request is received by the Slice Manager, until the Mobile Network is fully operational and ready to accept UE's connections. Slice Creation Time records are reported for each one of the Service Creation stages individually. The collected results are illustrated in

---

[3] http://literature.cdn.keysight.com/litweb/pdf/5989-5088EN.pdf http://literature.cdn.keysight.com/litweb/pdf/5989-5088EN.pdf

Figure 4-12. The results show that the average (excluding failed and outliers) time for service creation is 63.394 ms, the detailed results are provided in Annex Table A-22. The deployed service corresponds to the deployment of a vEPC instance accompanied by the configuration of the network and radio elements. It is expected that more complicated services may require more time, however this is also related to the VNF image size, virtualization infrastructure implementation and resources and the service forwarding graph complexity.

Failed service creation attempts recorded in 8% of the deployments due to a virtual machine (VM) boot time issue inside OpenStack. In these cases, boot times of the vEPC instances extended beyond the timers set by the Slice Manager and the NMS resulting to failed configuration attempts of the Mobile Network services. The penalty applied to the service creation time metrics in these cases is approximately 5 minutes plus the time it takes to re-instantiate the services. Plans for Phase 3 include the support for more types of core instances and the optimisation of the images stored inside the virtual infrastructure to minimize the failed attempts rate.

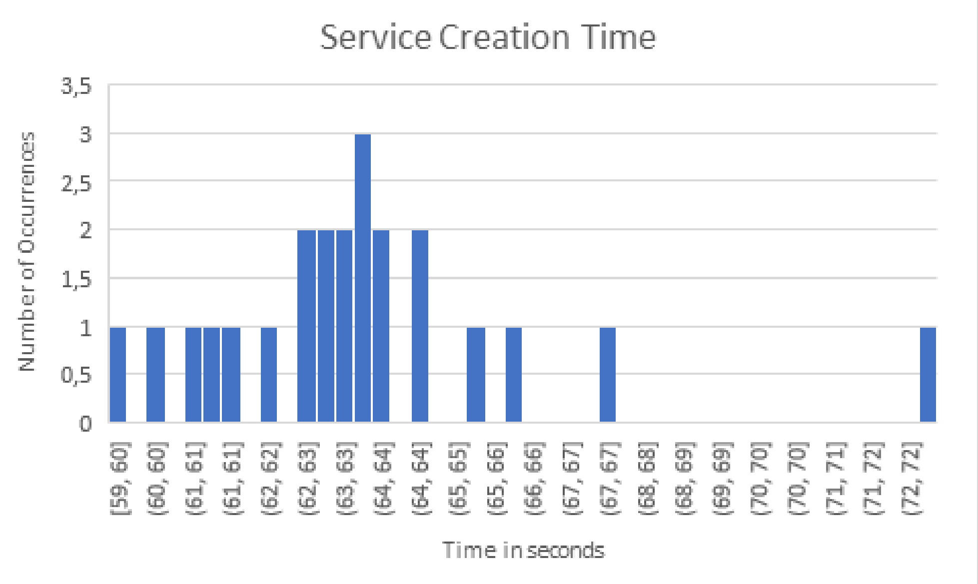


Figure 4-12 Service Creation Time Histogram for Athens Platform

4.2.2. Amarisoft RAN & Athonet CN (5G.3.Option3)

Figure 4-13 illustrates the test setup for the experiments we conducted in the Amarisoft RAN-Athonet CN 5G NSA setup in the Athens Platform. In all experiments, traffic flows between Endpoints 1 & 2, namely a Samsung A90 5G and a commodity Dell Laptop, respectively.

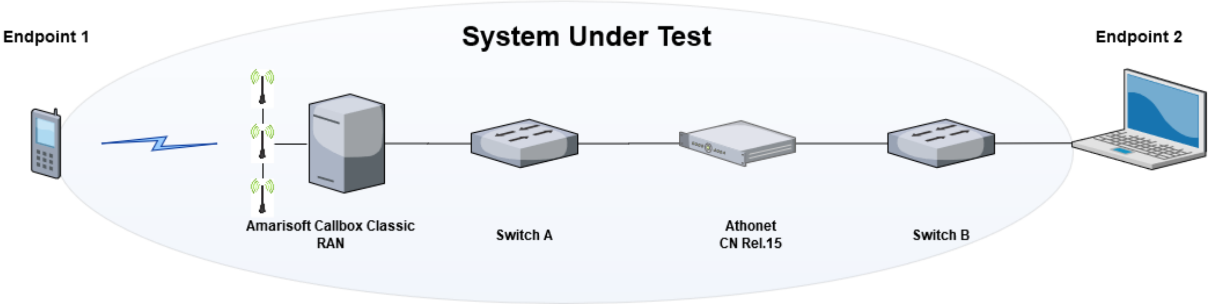


Figure 4-13 Amarisoft RAN – Athonet CN testbed setup

4.2.2.1. Throughput

During the experiment, the radio conditions were excellent and stable, as recorded in the UE by UMA's Resource Agent and are illustrated in Figure 4-14. The radio metrics captured were RSRP, RSRQ and RSSI with the following average values: RSSI= -51.00 +/- 0.00 dBm, RSRP= -71.00 +/- 0.00 dBm, RSRQ=-7.00 +/- 0.00 dB.

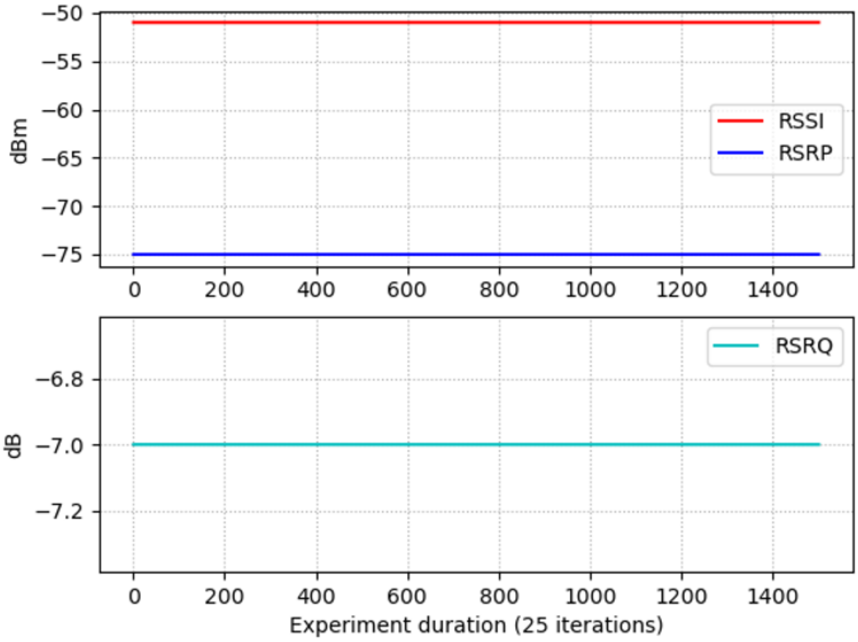


Figure 4-14 Radio conditions during Throughput experiments (Amarisoft RAN-Athonet CN)

The radio configuration of the 5G cell in this setup corresponded to 50 MHz Bandwidth, 2x2 MIMO, TDD, Band n78, 256 QAM DL, resulting in a theoretical throughput of 477 Mbps. As in the Amarisoft RAN-Amarisoft CN setup, we adjusted the maximum UDP data rate of iperf2 on 377 Mbps, providing a mean packet loss approximately around 1% during the experiment, as reported by the iperf2 probes.

The average Throughput in this case is 363.28 +/- 1.00 Mbps and the 95th percentile is 368.01 +/- 0.28 Mbps. Overall, this is slightly less than the reported throughput of the Amarisoft RAN-Amarisoft CN setup and this can be attributed in the transport network (10.2.1.0/16) between the Amarisoft RAN and Athonet CN. The Amarisoft RAN-Amarisoft CN setup does not have such transport network, as both domains communicate directly.

In this experiment, we notice again minimum values at 258.24 +/- 30.17 Mbps, which show up instantly among the data and can be attributed to filled up buffers from previous iterations. This is clearly shown by the 5th percentile result, where 5% of our measurements were below 362.75 +/- 2.41 Mbps. The detailed results are presented in Annex Table A-23

4.2.2.2. Round Trip Time

During the experiment, the radio conditions were excellent and stable, as recorded in the UE by UMA's Resource Agent. Figure 4-15 shows typical radio conditions throughout the experiments. The radio metrics captured were RSRP, RSRQ and RSSI and their detailed results are reported in the Annex Table A-24.

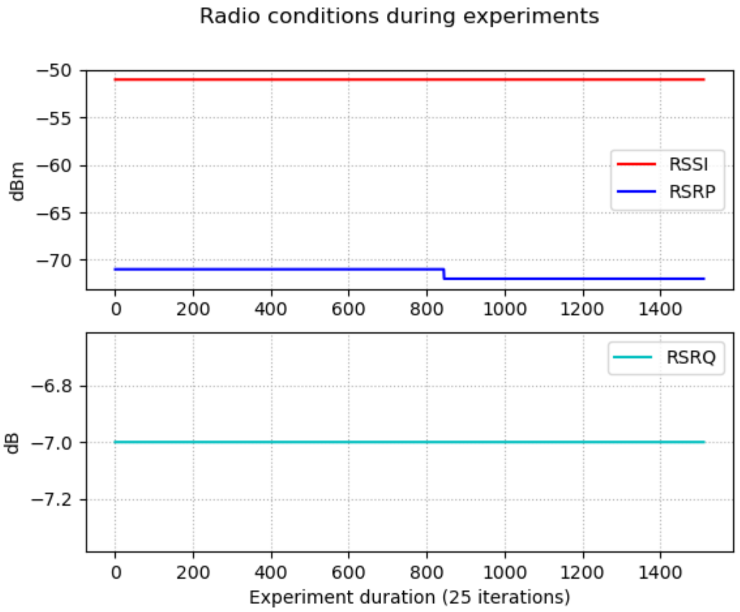


Figure 4-15 Radio conditions during E2E RTT per packet size experiments

The radio configuration of the 5G cell in this setup is 50 MHz Bandwidth, 2x2 MIMO, TDD, Band n78, 256 QAM DL, while the 4G cell provided 10 MHz Bandwidth, 2x2 MIMO, FDD in Band 1. The reported average E2E RTT of a single connected 5G COTS UE is 32.35 +/- 0.18 ms with 64 bytes packet size on an empty channel without background traffic. The 95th percentile E2E RTT is 41.42 +/- 0.45 ms. It is also important to note that all ping requests were successful, resulting in an average ping success ratio of 1.00, due to low network load and stable radio conditions throughout the experiment.

As expected, we show again in this setup that the RTT in 5G NSA networks is comparable to the 4G case, since the 5G NR Cell is anchored in an existing 4G deployment and uses the same CN (EN-DC). This clearly shows that initial 5G NSA deployments are suitable for supporting eMBB applications requiring higher throughput than the one 4G networks provide.

The detailed results of each packet size are provided in the Annex Table A-24. It is clear that packet size affects the E2E RTT, providing a range of average values from 32.39 +/- 0.21 ms (32 bytes) to 59.99 +/- 2.06 ms (1500 bytes). The bar chart of Figure 4-16 provides an overall overview of the percentiles of E2E RTT per packet size. It is important to note that E2E RTT for 32/64 bytes and 128/512 bytes almost overlap, indicating that the buffers of the SUT are optimized for traffic of such level. In addition, the standard deviation of E2E RTT is comparable for all packet sizes, ranging from 6.05 +/- 0.15 ms to 7.20 +/- 0.38 ms, indicating the stability of network conditions in the SUT.

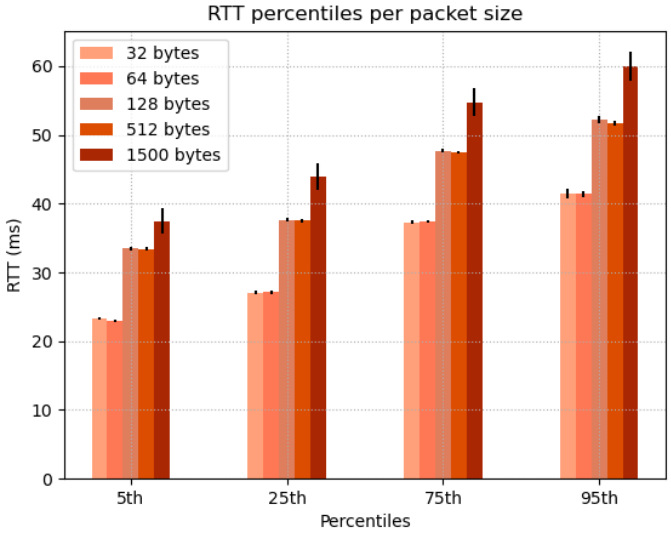


Figure 4-16 E2E RTT bar chart of percentiles per packet size in Amarisoft RAN-Athonet CN

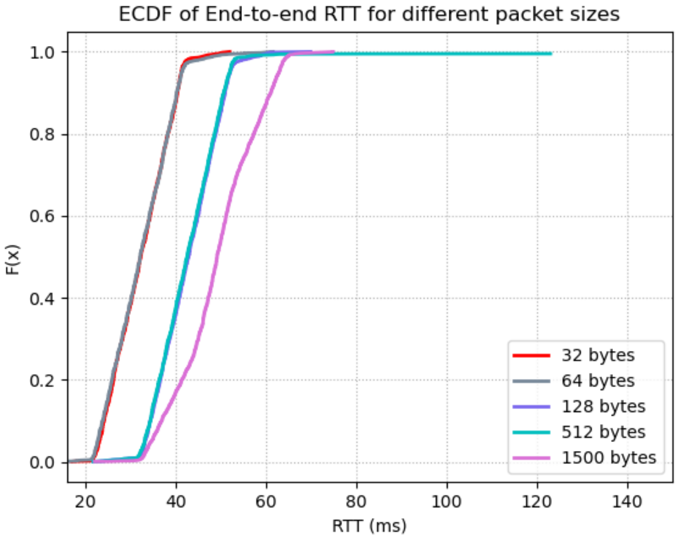


Figure 4-17 E2E RTT ECDF per packet size in Amarisoft RAN – Athonet CN

We also provide the ECDF of the data gathered in these experiments in Figure 4-17. The ECDF clearly presents the performance degradation as the packet size increases. We also notice here the E2E RTT overlap between 32-64 and 128-512 bytes in our SUT.

5. LIMASSOL PLATFORM EXPERIMENTS

5.1. Overview

The goal of the second phase of experimentation in the Limassol platform has been to:

- Verify the functionality of link aggregation across the dual (satellite and terrestrial) backhaul links;
- Validate the integration of the coordination layer and the test automation framework;
- Assess dynamic slice creation with VNFs at the core and the satellite edge;
- Measure 5G NR capabilities.

Table 5-1 lists the KPIs evaluated in the second trial and summarizes the kind of evaluation measurements conducted.

Table 5-1 Primary 5G KPIs evaluated at the Limassol Platform in the second phase

KPI to be evaluated at the Limassol Platform according to DoA	Evaluated in Phase 2	Comment
Ubiquity	No	Not scheduled for Phase 2
Latency	Yes	Phase 2 focused on RTT measurements (see below)
Reliability	No	Not scheduled for Phase 2
Service creation time	Yes	-
Additional 5G KPIs evaluated at the Limassol Platform	Evaluated in Phase 2	Comment
RTT	Yes	-
Throughput	Yes	-

Figure 5-1 depicts the physical topology of the Limassol platform, as it has been implemented for the Phase 2 experimentation campaign. This essentially corresponds to the configuration described in Deliverable D4.8 (The Limassol platform – Release B) [7], with a key addition – the RAN at the satellite edge has been upgraded to 5G NR, based on the Amarisoft Callbox solution[4]. However, due to the restrictions related to the COVID-19 pandemic, the RAN has not been yet integrated to the rest of the platform. That is, the 5G NR measurements were done in

---

[4] https://www.amarisoft.com/products/test-measurements/amari-lte-callbox/

a separate segment (with edge, EPC and RAN components) installed in the Space Hellas Cyprus (SHC) lab and currently detached from the rest of the platform.

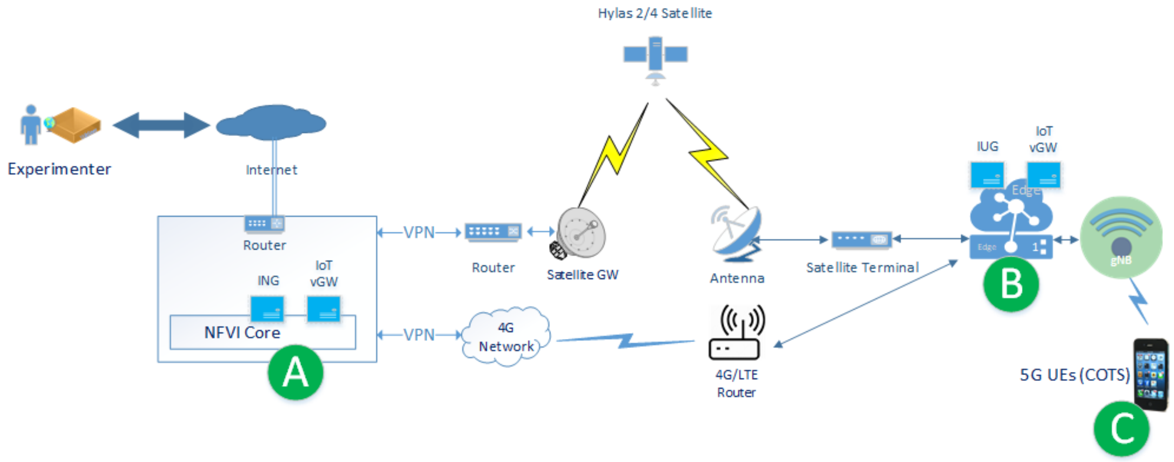


Figure 5-1. Actual topology of Limassol platform implemented for Phase 2 experimentation and measurement points

The measurement campaign described in the next section focuses on assessing the key features added during Phase 2, namely:

- The Coordination Layer, along with the Slice Manager, deployed as VMs in the core (OpenStack-based) infrastructure.
- The link aggregation virtual functions, namely the Intelligent Network Gateway (ING) deployed at the core, and the Intelligent User Gateway (IUG) deployed at the satellite edge.
- The terrestrial backhaul, implemented using 4G links.
- The 5G RAN at the satellite edge.

Figure 5-1 also displays the main reference points used for the measurements. All tests were carried out between (or at) these points.

- Reference point A: At the platform core compute infrastructure.
- Reference point B: At the satellite edge compute infrastructure.
- Reference point C: At the 5G UE.

5.2. Experiments and Results

5.2.1. Edge site – Core data-centre measurements

Traffic between core datacentre and edge site is handled from the multilink mechanism that equally leverages both satellite and LTE backhaul for better performance or routes all traffic through one link if the other one is unavailable. The following measurements include all possible scenarios such as both links and one link at a time. Since the split logic is performed on the core datacentre, throughput measurements concentrate on downlink traffic only.

5.2.1.1. Downlink throughput (goodput) – satellite backhaul only

This test measures the downlink throughput between the CN (point A) and the edge node (point B), when the terrestrial backhaul is unavailable.

Figure 5-2 depicts the goodput (in Mbps) measured on the satellite link, i.e. when the terrestrial link is unavailable. The iPerf tests have been repeated 25 times for a duration of 180 seconds for each iteration. We have used boxplots as we want to focus on the variability of the obtained goodput. The results show that the goodput is about 10.7 Mbps for all the iterations. Furthermore, we note that the 95th and the 5th percentiles of each boxplot are close that they overlap. As the traffic is sent only over one link, so no overhead has been generated, which explains the stability of the goodput.

Detailed results are provided in Annex Table A-25.

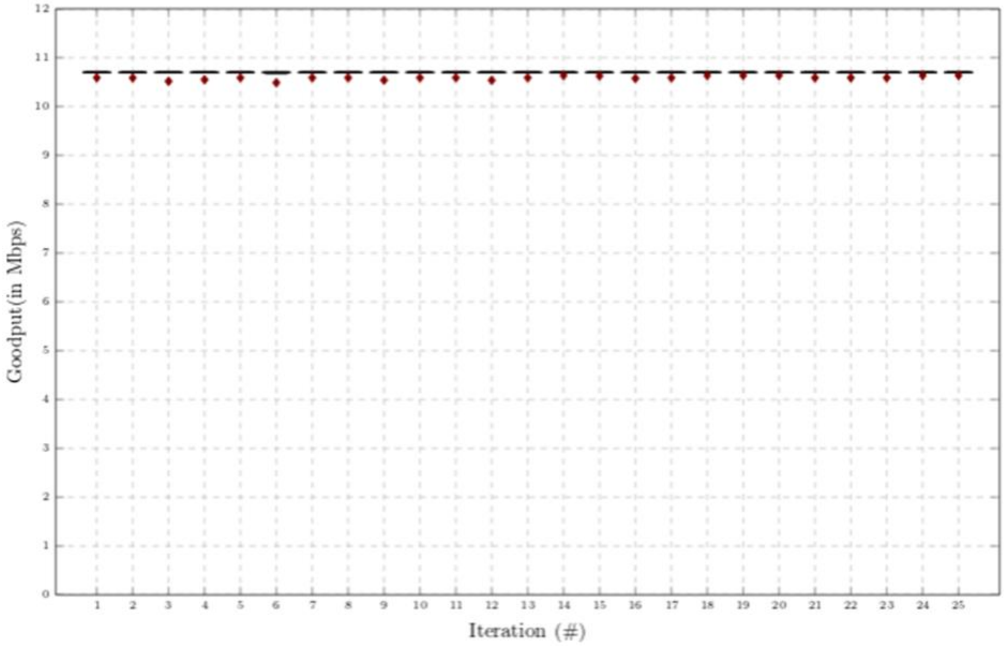


Figure 5-2. Downlink throughput – satellite backhaul only

5.2.1.2. Downlink throughput (goodput) – terrestrial backhaul only

This test measures the downlink throughput between the core network (point A) and the edge node (point B), when the satellite backhaul is unavailable.

The iPerf traffic originates from the core and traverses the two VNFs ING (at the core network) and IUG (at the edge network). However, since the satellite backhaul is not available, having been manually disabled (emulating conditions where the edge segment is outside satellite network coverage), the entire traffic is sent over the terrestrial backhaul. The goodput results are shown in the following.

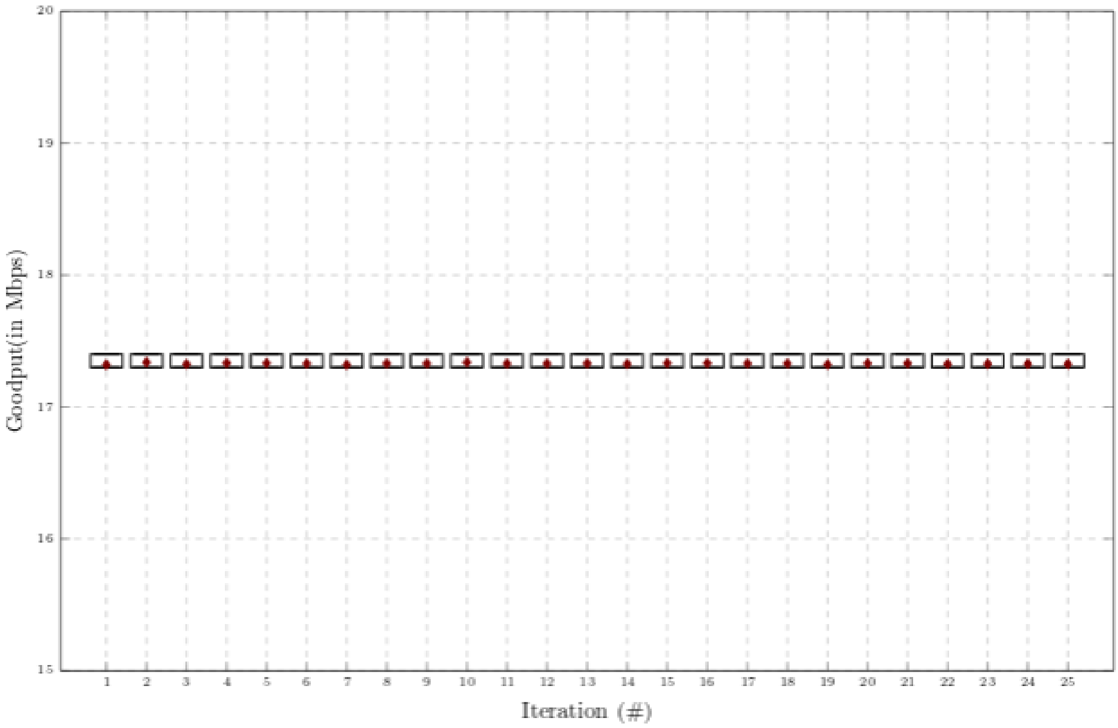


Figure 5-3. Downlink goodput – terrestrial backhaul only

Figure 5-4 shows the goodput (in Mbps) measured on the terrestrial link when the satellite link is down. We have done the same tests as for the previous tests. The terrestrial link is 20 Mbps bandwidth and 100 ms RTT. The obtained results turn around 17.3 Mbps, coherent with what is seen on the satellite test.

Detailed results are provided in Annex Table A-27.

5.2.1.3. Downlink throughput – link aggregation

The traffic originates from the core and traverses the virtual link aggregation functions (ING and IUG) at the core and the edge. In this case, both backhauls are enabled, emulating conditions where the edge segment is within terrestrial coverage. The MPTCP link bonding feature enables the use of the terrestrial link in order to boost the overall throughput, even for single-flow connections, which are split across the two links.

We have used a Weighted Round Robin (WRR) policy with the weights 5 and 3. That is, the traffic will be split accordingly; 5 packets will be sent on the terrestrial link, and 3 packets on the satellite link.

The results show that the per-user throughput benefits from the total combined terrestrial 17.3 Mbps and satellite 10.7 Mbps goodput (avg. 27.7 Mbps), thanks to the link bonding achieved by the link aggregation MPTCP VNFs deployed at the core and edge.

Detailed results are provided in Annex Table A-27.

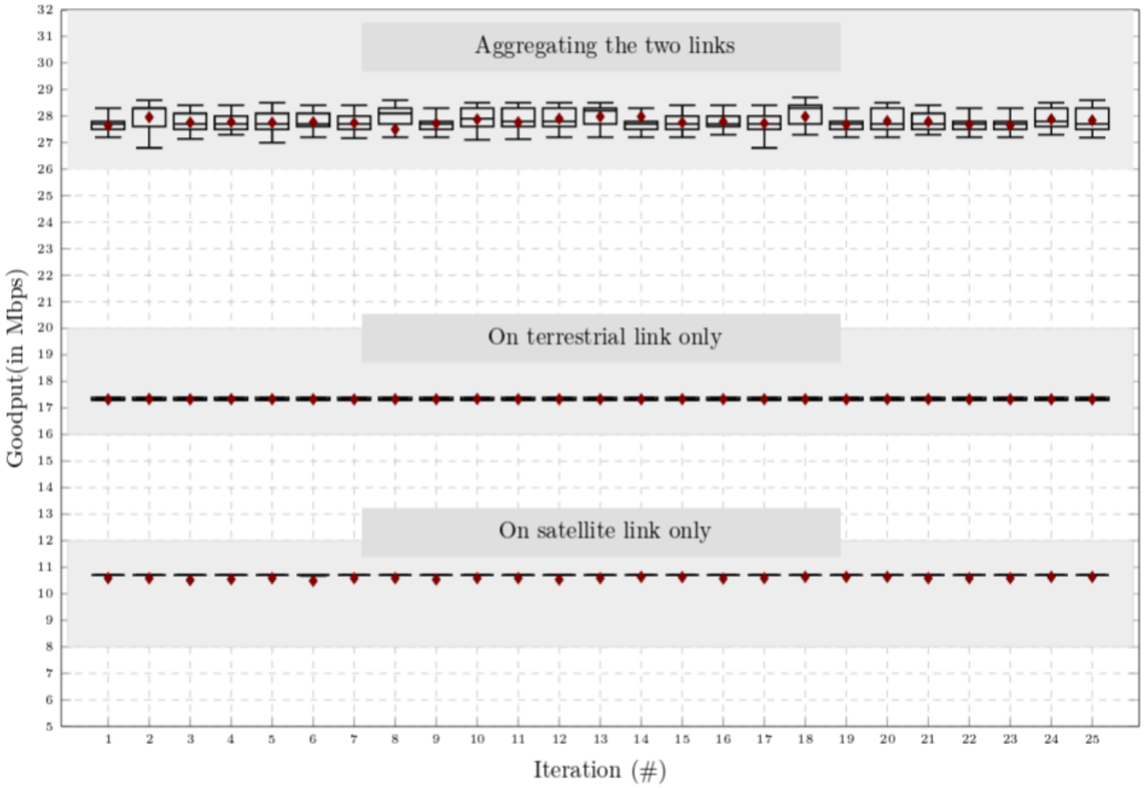


Figure 5-4 Downlink throughput – link aggregation

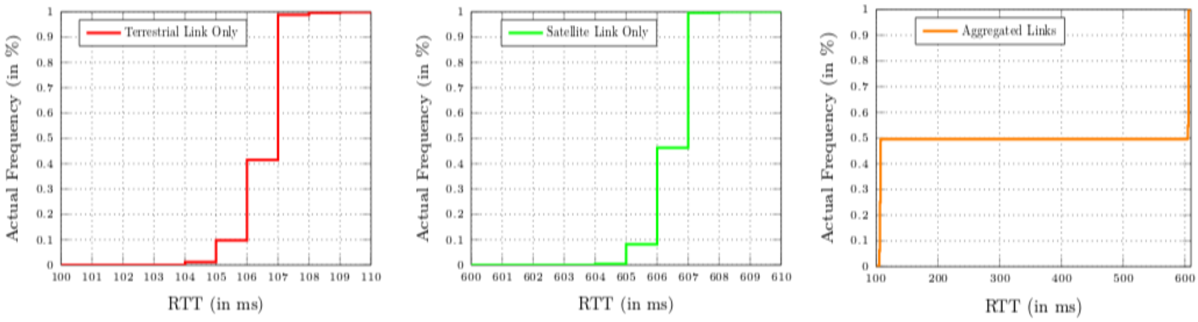
5.2.1.4. RTT – satellite and terrestrial backhauls and link aggregation

We have performed pings between two Linux VMs, one located on the edge and the other one on the core. The pings cross the VNFs IUG and ING. The results are shown in Figure 5-5. The graph depicts the ratio of ICMP packets in a population of 1000 packets that are generated in less than x ms.

Figure (a) shows the ratios when we use only the terrestrial backhaul (i.e. the satellite backhaul is down). In this figure we can notice that 40 % of the ICMP packets have been received in less than 106 ms, while all the sample has been received after 108 ms.

Figure (b) presents the ratios for the satellite backhaul only (i.e. the terrestrial backhaul is down). Here, around half of the ICMP packets have been received in 606 ms, and the entire packets have been received one second latter.

Figure (c) depicts the ratios when the two backhauls (ie, terrestrial and satellite backhauls) are used. In this figure we want to show how the ICMP packets have been achieved through the two backhauls. Comparing to figures (a) and (b), in figure (c), we notice that half of the traffic is received in less than 108 ms, whilst all the traffic was received after 607 ms. This means that the average latency when using both backhauls is equal to half of the sum of the two latencies for the terrestrial and satellite backhauls (i.e. 357.5 ms).



(a) Ping over terrestrial link only (b) Ping over terrestrial link only (c) Ping over the two links

Figure 5-5 CDF for the measured RTT on the two links (Satellite and terrestrial)

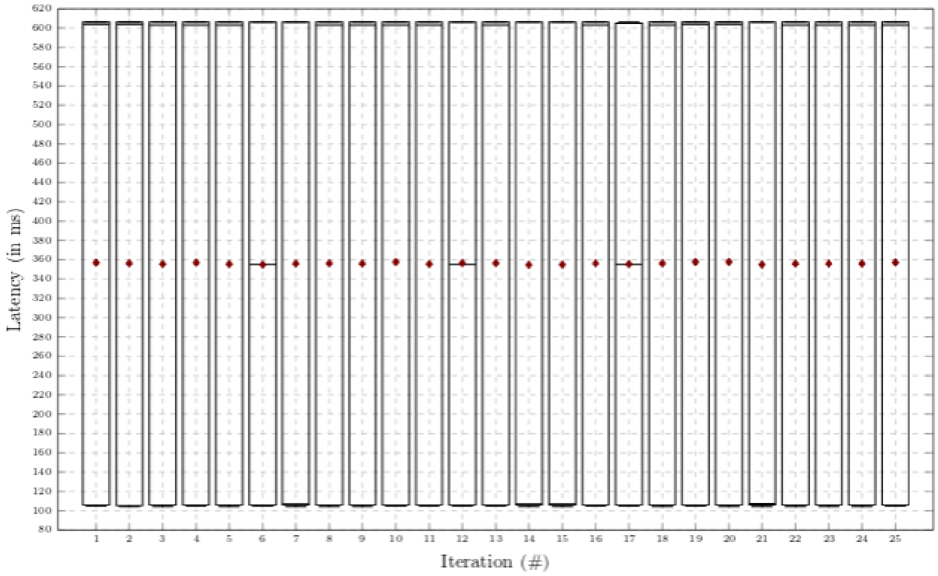


Figure 5-6 RTT - Link aggregation

The results in Figure 5-6 essentially show the behaviour of the link aggregation VNFs. Since the average RTT over satellite is 607 ms, while over the terrestrial is 108 ms, the end results yield a mean value of 357.5 ms, with a substantial standard deviation (250 ms).

The results are detailed in Annex Table A-28.

For future research directions, it would be of interest to experiment other policies such as Path Selection-Based on Object Length (PSBOL) and Offload, to stabilise and optimise the latency.

5.2.1.5. Slice creation time - Core DC

This test measures the creation time of a slice consisting of a VNF located at the core DC (point A). The test case, initiated from the Portal and ELCM, is passed to the Slice Manager (Katana), which requests from OSM to create/ delete the VNF and gathers the results.

The average service creation time (essentially the VNF spin-up time) is around 45 secs. It can be also seen in Figure 5-7 that this time is more or less constant across iterations. Naturally, it depends on the performance of the OpenStack cluster and –more importantly- on the size of the VNF instance. The detailed results are provided in the Annex Table A-29.

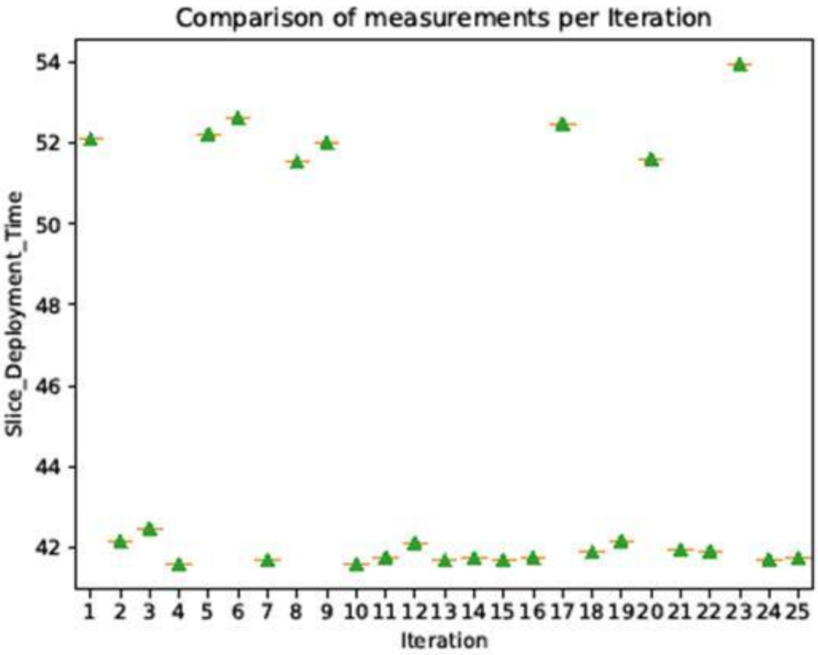


Figure 5-7. Slice creation time – Core DC

5.2.1.6. Slice creation time - Edge node

This test measures the creation time of a slice consisting of a VNF located at the edge node (point B). Similarly, the test case, initiated from the Portal and ELCM, is passed to the slice manager (Katana) which requests from OSM to create/ delete the VNF and gathers the results.

The average service creation time (essentially the VNF spin-up time) is around 102 secs. Apart from a couple of outliers (see Figure 5-8), it can be also seen that this time is more or less constant across iterations. The detailed results are provided in Annex Table A-30.

It is obvious that the average service creation time at the satellite edge is almost doubled, compared to the core (discussed in the previous section). This is due to the lower performance of the edge node (compared to the core DC servers) as well as the delay of the satellite link, which slows down the exchange of control traffic.

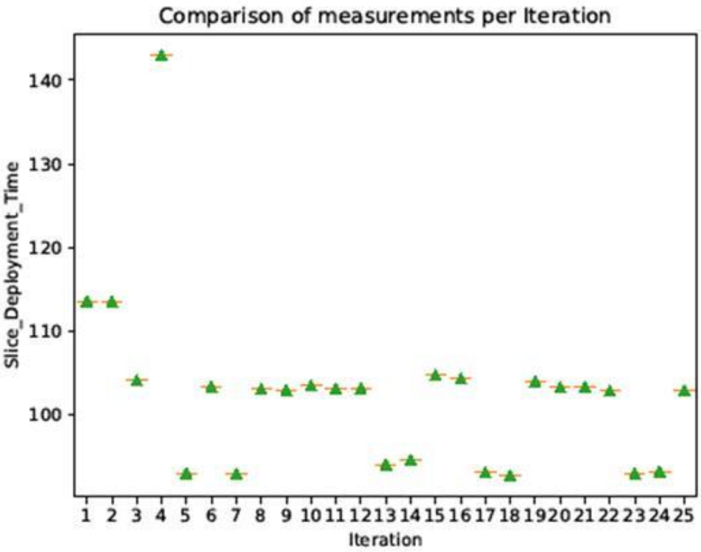


Figure 5-8. Slice creation time – Edge DC

5.2.2. 5G RAN measurements

This section concentrates on the performance of the 5G RAN. Measurements are made between the edge node (point B) and the UE (point C). The 5G RAN operates in Band n78 (3.5GHz), 50 MHz bandwidth, 2x2 MIMO configuration, non-stand-alone (NSA) mode. The tests are done in lab, over-the-air. The UE used is a Samsung Galaxy A90 5G smartphone.

5.2.2.1. Downlink throughput

The throughput is measured using iperf3 agents at the edge node and the UE, in ideal line-of-sight conditions. The distance between the gNB and the UE is approx. 2m.

The DL throughput measurement is on average 210 Mbps, close to the one theoretically expected for the given RAN configuration. The network behavior was almost constant; the standard deviation between the measurements was relatively small.

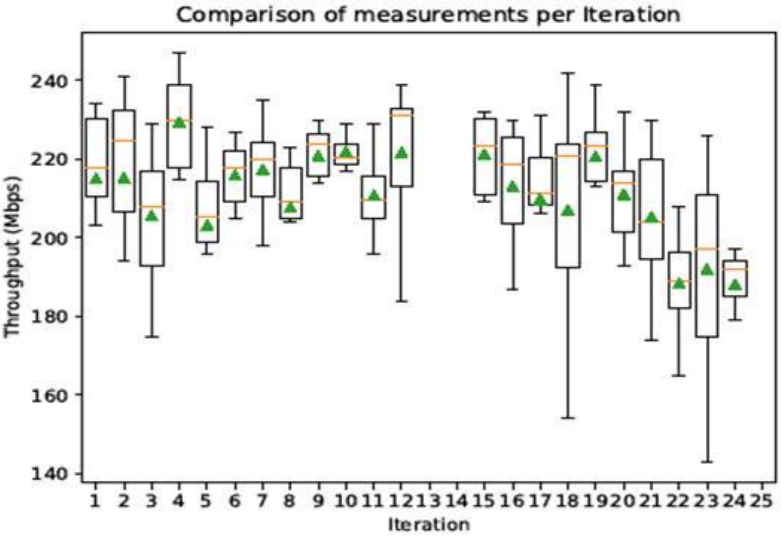


Figure 5-9. Downlink throughput – 5G RAN

Detailed results are provided in Annex Table A-31.

5.2.2.2. Uplink throughput

The throughput is measured using iperf3 agents at the edge node and the UE, in ideal line-of-sight conditions. The distance between the gNB and the UE is approx. 2m. The UL throughput measurement is on average 40 Mbps, close to the one theoretically expected for the given RAN configuration. The network behavior was almost constant, as it is shown in Figure 5-10. The standard deviation between the measurements was minor.

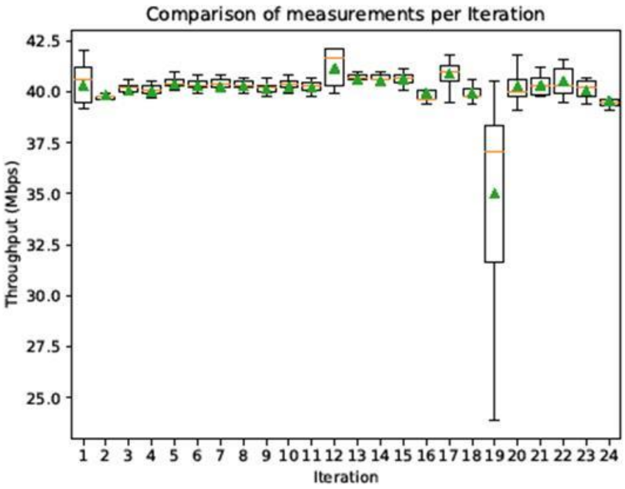


Figure 5-10 Uplink throughput – 5G RAN

Detailed results are provided in Annex Table A-32

5.2.2.3. RTT

The RTT is measured using ping agent at the edge node, in ideal line-of-sight conditions. The distance between the gNB and the UE is approx. 2m. The RTT measured was 32 msec on average, almost constant across measurements. Details are provided in Annex Table A-33. The RTT test shows no significant improvement over the LTE RTT measured in deliverable D6.1. We expect such improvement when we switch from NSA to SA mode, in Phase 3 measurements.

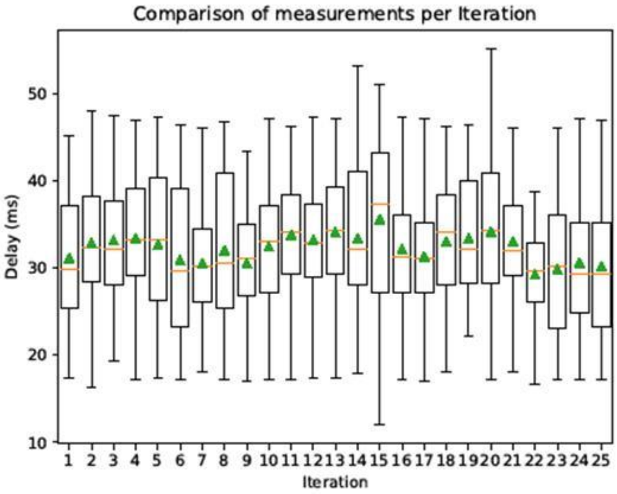


Figure 5-11. RTT – 5G RAN

6. SURREY PLATFORM EXPERIMENTS

6.1. Overview

This section provides the results of the experiments conducted in the Surrey Platform. The first set of experiments consists of the baseline Throughput and RTT tests, conducted in the context of the 5GNR Rel.15 NSA network. The second set of experiments comprise a number of unitary tests that aim at verifying the basic E2E reliability and performance of the Surrey Platform IoT use case, involving Wi-Fi and LoRA radio channels. Finally, results of tests on NB-IoT Energy Consumption and Coverage are also provided and discussed.

6.2. Experiments and Results

6.2.1. 5G NR (NSA) [Rel. 15]

For the purpose of measuring the RTT and throughput in the Surrey Platform, the network setup consists of the following components:

- **Core**: Rel.15 4G Core NSA
- **Control Plane**: 4G RAN
- **User Plane**: 5G RAN (Huawei Commercial)
- **UE**: 5G CPE

The setup considered is depicted in Figure 6-1:

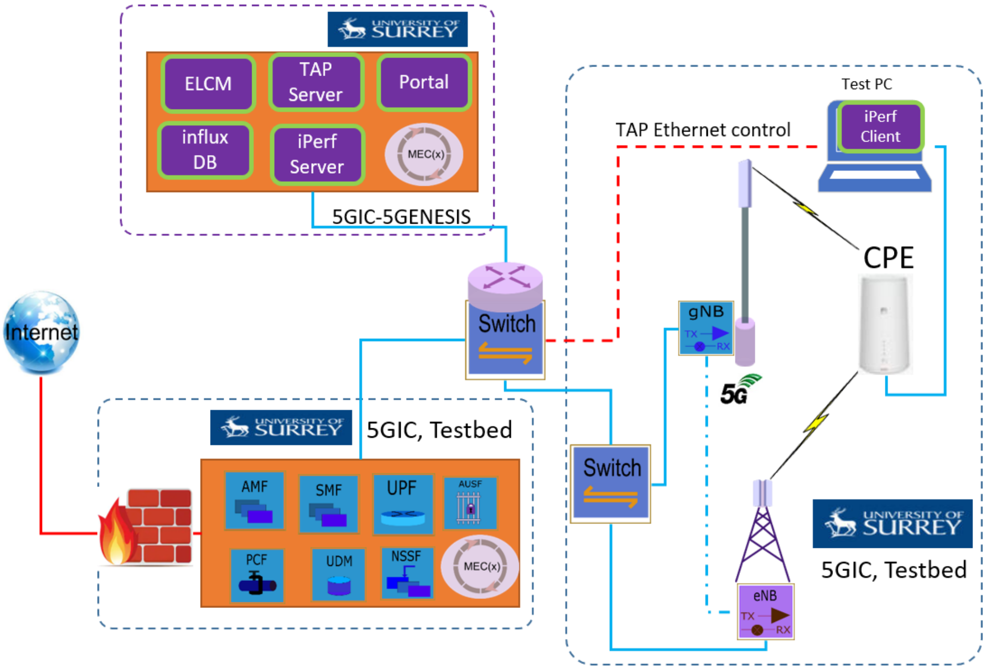


Figure 6-1. Architecture of the 5GNR network used for the Surrey platform tests

6.2.1.1. Round Trip Time

The purpose of this test case (*TC_RTT_e2e*) is to assess the end-to-end RTT from a 5G Customer Premises Equipment (CPE) client to a server over a 5GNR NSA Rel.15 mobile network.

The 95th percentile RTT is 12.54 +/- 0.05 ms. It needs to be noted that the measured RTT values are lower than the respective values in similar setups of the other 5GENESIS platforms. This is a result of the fact that the 5G RAN components in the Surrey Platform are connected to the SDN switch in the core network via fast and reliable fibre links. Moreover, the Surrey Platform 4G network consists of powerful fast performing servers that are able of performing all required tasks in a computationally efficient manner, thus reducing the resulting latencies.

The results of the round-trip-time test are shown in Figure 6-2. Detailed results of this test case (including 5th, 25th, and 75th percentile RTT) can be found in Annex Table A-37.

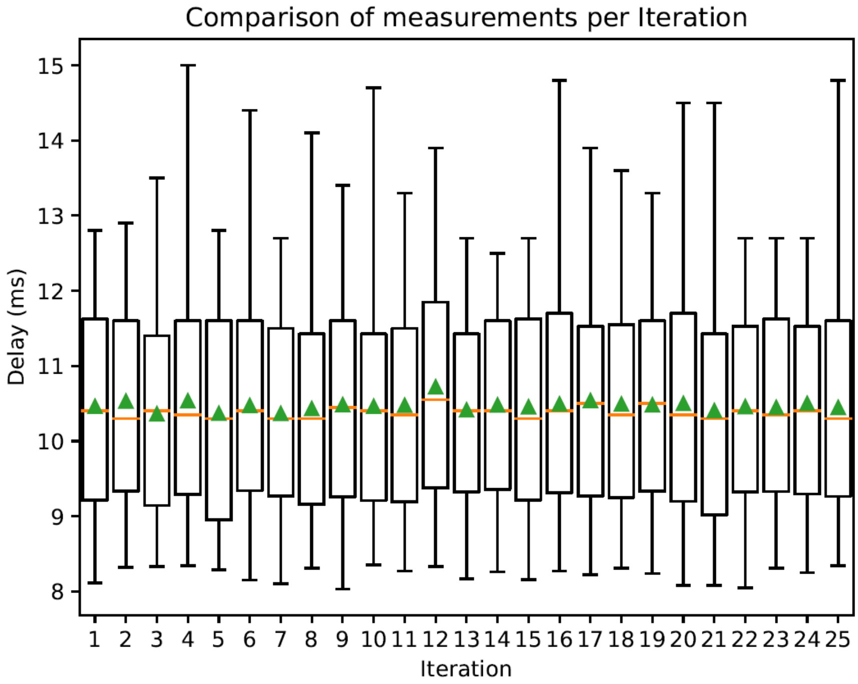


Figure 6-2. Round Trip Time test results

6.2.1.2. Throughput

This section presents and discusses on the results of the Throughput tests performed in the UL and DL directions of the 5GNR Rel.15 NSA configuration in the Surrey Platform.

In the UL direction, the throughput of both UDP and TCP protocols was measured. For UDP, the 95th percentile throughput was 122.54 +/- 0.36 Mbps (see TC-THR-Udp), while for TCP the respective throughout value was 67.52 +/- 5.04 Mbps (see TC-THR-Tcp).

The results of the UL throughput tests are illustrated in Figure 6-3 (UDP) and Figure 6-4 (TCP). In the case of the TCP throughput, it can be seen that there is a somehow increased variance of the results. This can be mainly attributed to environmental factors at the time of testing, since the Surrey Platform consists of and outdoors testbed.

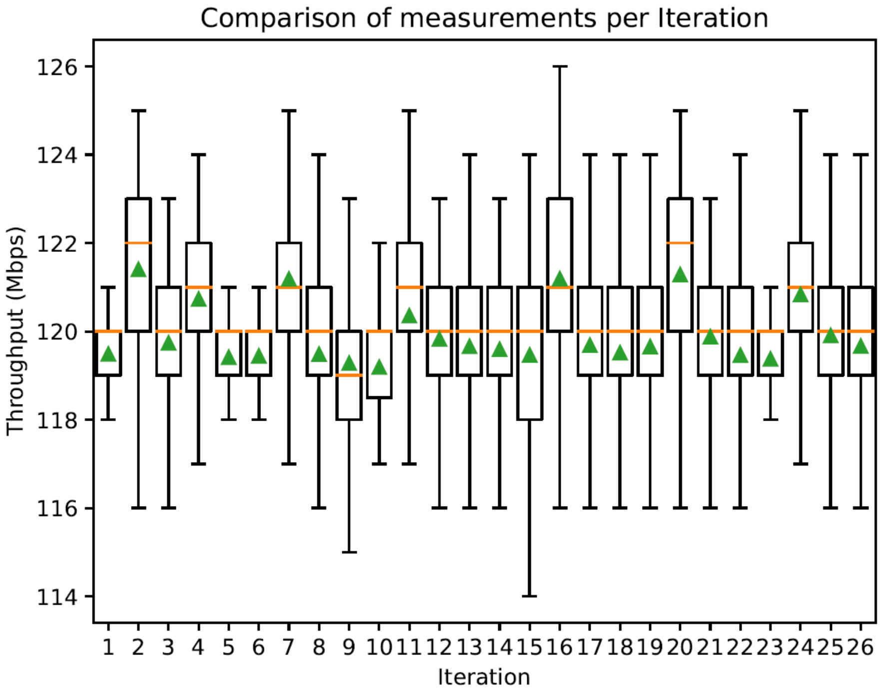


Figure 6-3. Throughput test results (Uplink, UDP)

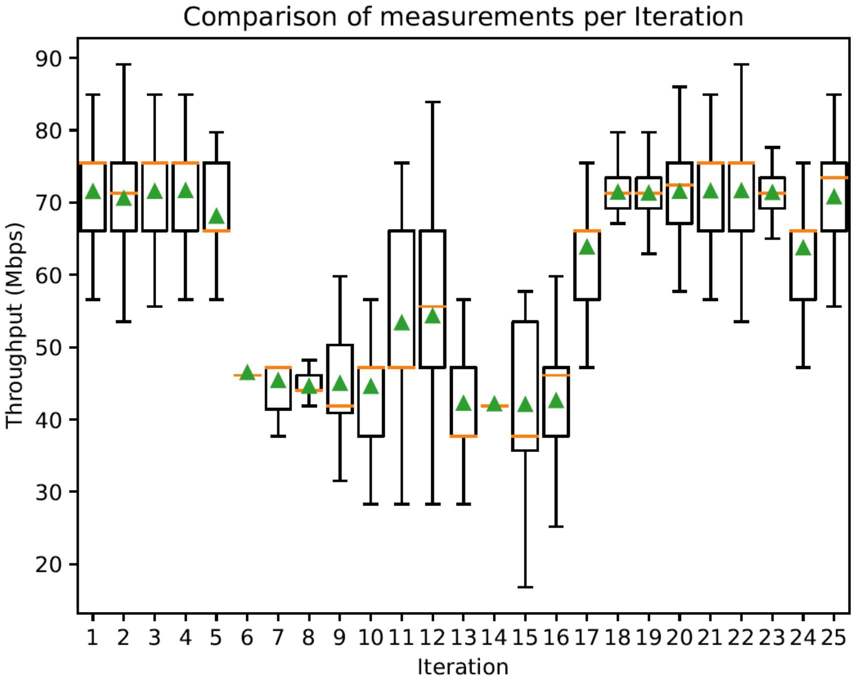


Figure 6-4. Throughput test results (Uplink, TCP)

In the DL direction, the 95th percentile throughput is 651.83 +/- 10.30 Mbps (see TC-THR-Tcp for more detailed results). The results of the DL throughput test are also depicted in Figure 6-5.

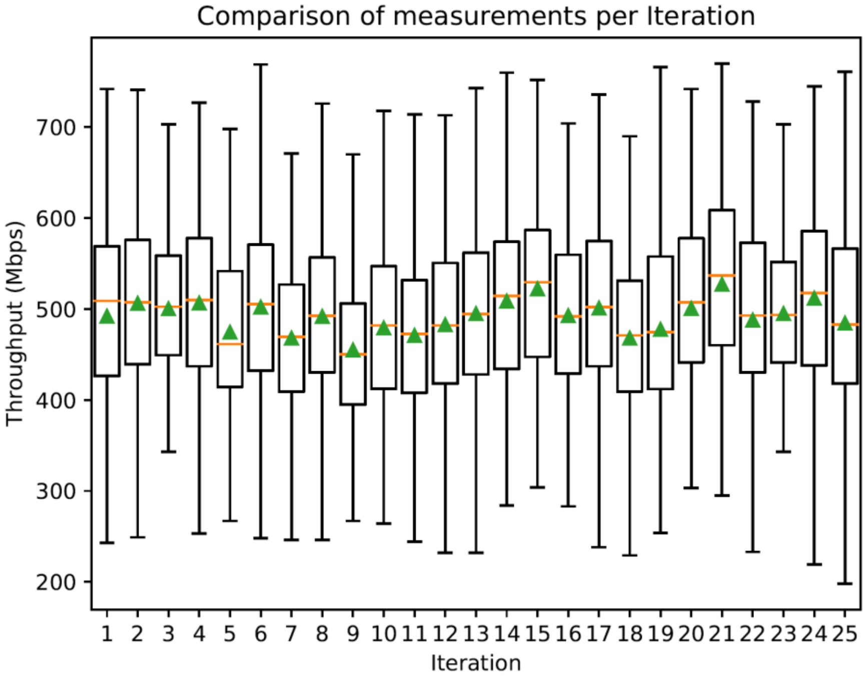


Figure 6-5. Throughput test results (DL. TCP)

Detailed results are provided in the Annex (see Annex Table A-34, Annex Table A-35 and Annex Table A-36).

6.2.2. IoT use case experiments

As a quick reminder, the IoT experiment is based on Pycom/Pysense boards. Pycom boards provide support for various radio channels including in particular LoRA and Wi-Fi, while Pysense provides Pycom boards with a shield supporting sensor reading. The targeted deployment of Pycom/Pysense boards is about 300 units. This experiment consists of reading sensor values and sending them –in JSON format– relying on various protocols, namely HTTP, MQTT and the Constrained Application Protocol (CoAP), all over Wi-Fi[5].

As far as LoRA is concerned, sensor data is collected and sent to the The Things Network (TTN) network as an array of hexadecimal values, before being converted to the normal JSON format using a JavaScript conversion function (uploaded to and applied at TTN website). Then the JSON data is published by TTN along specific topics which in turn the MQTT client at the platform side subscribes to.

This first series of test does not rely on full Pycom boards deployment due to technical and logistic difficulties resulting from the COVID-19 pandemic (e.g. inability to access offices and labs, last order of Pycom boards put on hold at the supplier side).

These unitary tests will provide basic E2E reliability and performance test involving both Wi-Fi and LoRA radio channels. As far as Wi-Fi is concerned, we will test the HTTP POST protocol only.

Two different tests are conducted in this section:

---

[5] : only Wi-Fi is used for CoAP, MQTT and HTTP POST as it happens that, unfortunately, the Pycom/Pysense boards LTE/NB-IoT do not support the two LTE bands allocated to the 5GIC testbed.

- **Performance test:** for both LoRA and Wi-Fi, we send 4 batches of IoT packets (respectively 1, 10, 100 and 1000 packet(s)). Then for each batch we calculate the average E2E delay between time at packet emission (timestamp embedded within the payload) and time at packet storage within the mySQL database. Both clocks at the board and server sides are synchronized to a NTP server (e.g. time.google.com). We do expect significant differences between HTTP and LoRA-based performance tests as LoRA involves a 3rd party in the middle (TTN). As far as Wi-Fi is concerned, we conduct the tests with the three protocols (HTTP POST, MQTT and CoAP) with various sampling rates as will be indicated in the Test Results section;
- **Reliability test**: for both LoRA and Wi-Fi: during the performance test described above we do a secondary measurement about the number of received IoT packets and assess the packet loss rate.

6.2.2.1. Packet Delay and Packet Loss

Before going in detail into the results of the unitary tests, we remind what we mean by sampling rate. This rate (expressed in seconds) is an additional time we add at the end of each of the data collecting/sending cycle, which takes at most 1.1 s to perform. Increasing this value decreases the frequency of such cycles.

When conducting the tests for each protocol used with Wi-Fi we start with a sampling rate of 0, meaning we try to send data as much as possible, i.e. every 1.1 s in average. Whenever the results are not satisfactory, we add some extra time in order to see if the protocol under test performs better with this extra time and therefore under lesser workload.

It is worth noting that, as far as LoRA is concerned, sending packets at the highest rate results always in packets being rejected, due to a breach to the European regulations applied to LoRA (see the discussion at the bottom of Annex Table A-41.

The results obtained from the tests (see Annex Table A-38, Annex Table A-39, Annex Table A-40 and Annex Table A-41) verify the correct implementation and integration of the system E2E. In the tests regarding HTTP and MQTT, the results indicated 100% success ration in the received packets while, in the test regarding CoAP, the success ratio is 53.9%, leaving room for improvement in near real-time packet transport scenarios. As far as packet delay is concerned the delays are minimal and of the same scale (92 ms-143 ms) in all three tests.

6.2.2.2. Narrowband IoT Energy Consumption

Narrowband IoT (NB-IoT) is a cellular Low Power Wide Area (LPWA) technology that was first introduced in 3GPP Release 13. NB-IoT devices are expected to operate unattended, potentially in inaccessible and signal-challenged locations, for at least 10 years on a single battery charge (~5 Wh). In view of this, it is essential to verify that an NB-IoT device is indeed able to operate during a 10-year period, and to understand how the energy consumption is affected by factors such as the traffic intensity and burstiness, and by different tunings of the NB-IoT stack, including power saving mechanisms such as discontinuous reception (DRX), extended idle mode DRX (eDRX), and Power Saving Mode (PSM).

We studied and verified the energy consumption of an NB-IoT device when data traffic is routed through the control plane and uses the control-plane CIoT EPS optimization procedures. All tests were carried out through simulations in the OMNeT++ simulation environment. To accomplish this, we have implemented a simulation model depicted in Figure 6-6, which includes all key network elements involved in an uplink transfer between an NB-IoT device and an application server.

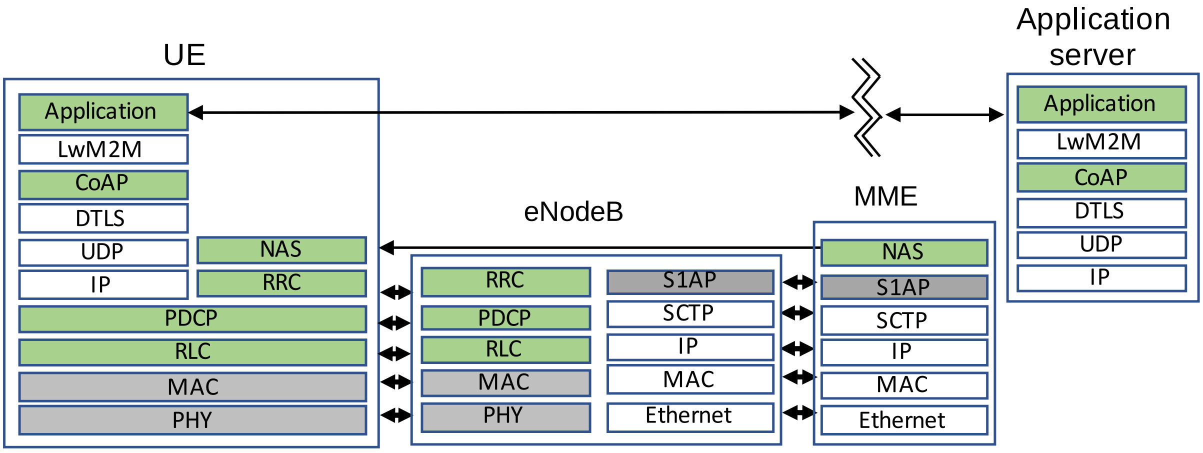


Figure 6-6 The modeled NB-IoT network architecture. The green-colored protocol layers are modeled in detail; the grey-colored, and even more so the white-colored layers, only model the essential parts of their functionality.

The target KPI used for evaluation is the average energy consumption of an NB-IoT device over a ten-year period of operation (NB-IoT Device Energy Consumption). In our tests, we considered a scenario with good radio coverage, and, in view of this, made the power consumption assumptions in Table 6-1.

Table 6-1 Power consumption assumptions

Activity	Power Consumption (mW)
Transmit	545
Receive	90
RRC Idle State (light sleep)	3
PSM mode (deep sleep)	0.015

6.2.2.3. Device Energy Consumption

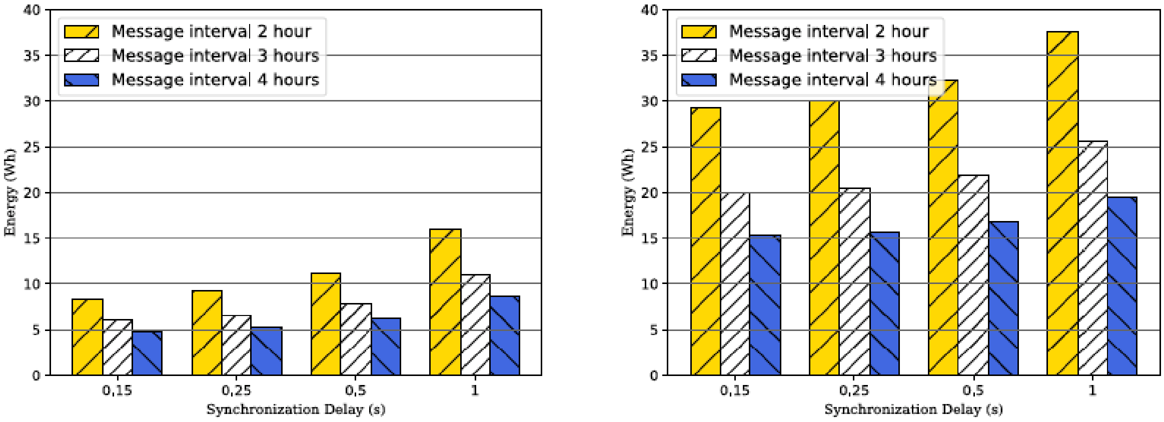
In our tests, an NB-IoT device transmitted messages with different periodicity and in bursts of different lengths. The transmission periods and burst sizes considered are listed in Table 6-2.

Table 6-2 Traffic properties of NB-IoT traffic

Traffic Property	Values
Length of transmission period (hours)	2, 3, 4
Burst length (# messages)	1, 2, 3

Each test was repeated 30 times, and the average NB-IoT Device Energy Consumption was used as an estimate of the expected amount of energy consumed during a 10-year period. Annex Table A-42 details the test case for the "NB-IoT Device Energy Consumption" test.

Annex Table A-43 illustrates the large impact C-DRX has on the average NB-IoT device consumption in scenarios with different network synchronization delays. In these scenarios, the NB-IoT device stack was configured with an Inactivity timer of 20 s, an Active timer of 30 s, and a CoAP retransmission timer of 2 s. Uplink transfers took place with a burst length of 1 message. As follows, our tests suggest that C-DRX is essential to keep the average NB-IoT Device Energy Consumption below 5 Wh. To this end, in the remainder of this section, we only consider the tests conducted with C-DRX enabled.



C-DRX Enabled C-DRX Disabled

Figure 6-7 The effect of C-DRX on energy consumption

It also follows from Figure 6-7 that the periodicity had a significant effect on the average NB-IoT Device Energy Consumption, however, less so than C-DRX. Still, we observe that in the simulations considered, the average NB-IoT Device Energy Consumption was only below 5 Wh in those tests where the synchronization delay was short. Moreover, it could be noted that the effect of the periodicity increased when the transmission quality of the air interface deteriorated, and the synchronization delay increased.

The impact of the message burst length on the average NB-IoT device consumption in scenarios with the Inactivity, Active, and CoAP retransmission timers configured as before; and, with uplink transfers generated every third hour is shown in Figure 6-8. Again, we studied how the energy consumption varies with increasingly more challenging air interface conditions, i.e., with longer synchronization delays.

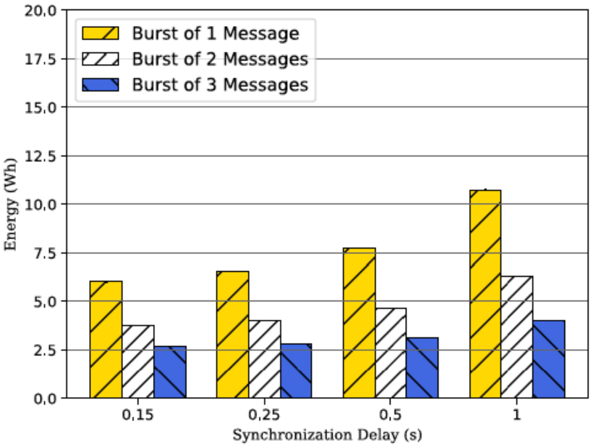


Figure 6-8 The effect of message burst length on energy consumption

It clearly follows from these tests that the length of the message bursts indeed seem to have a significant impact on the average NB-IoT Device Energy Consumption, and that the effect increases with increasing synchronization delays. In fact, in these tests, grouping several messages together was necessary to keep the average NB-IoT Device Energy Consumption below 5 Wh.

The results are detailed in Annex Table A-42, Annex Table A-43, Annex Table A-44 and Annex Table A-45.

6.2.2.4. Narrowband IoT Coverage

As mentioned in the previous section, NB-IoT devices operate in heterogeneous environments, such as deep indoor, indoor, and outdoor scenarios. Hence, it is important to empirically analyze the availability of the NB-IoT signal at different locations, in order to identify correlations and causalities between Radio Access Network (RAN) deployment and coverage performance, ultimately moving toward system improvement.

We thus conducted a large-scale measurement campaign of NB-IoT coverage in the city of Oslo, Norway, covering three weeks during summer 2019. We used the Rohde&Schwarz (R&S) TSMA6 toolkit, which includes a spectrum scanner and a software for data collection and visualization named ROMES4, together with an Exelonix Narrowband (NB) USB device and a global positioning system (GPS) antenna, as shown in Figure 6-9:

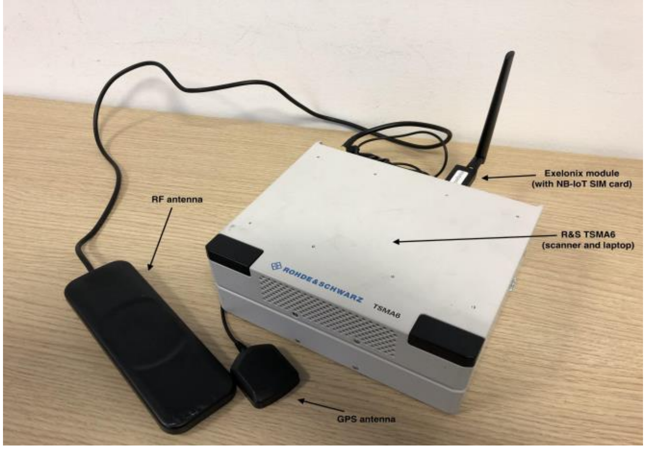
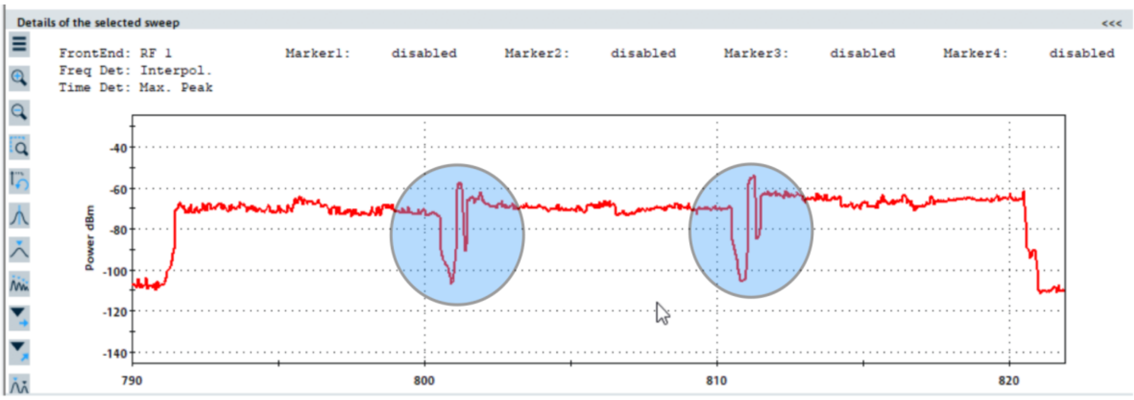


Figure 6-9 Measurement setup for NB-IoT RAN coverage measurement in Oslo, Norway



Details of the selected sweep

```
FrontEnd: RF 1            Marker1:   disabled    Marker2:   disabled    Marker3:   disabled    Marker4:   disabled
Freq Det: Interpol.
Time Det: Max. Peak
```

Figure 6-10. NB-IoT carriers in LTE Bard 20 (guard bands) for two Norwegian operators.

We enabled the scanner to perform passive measurements on four LTE bands (including guard bands), i.e., Band 1, 3, 7, and 20. We detected three LTE operators; two of them also provide NB-IoT carriers in the guard bands of LTE Band 20, as shown in Figure 6-10.

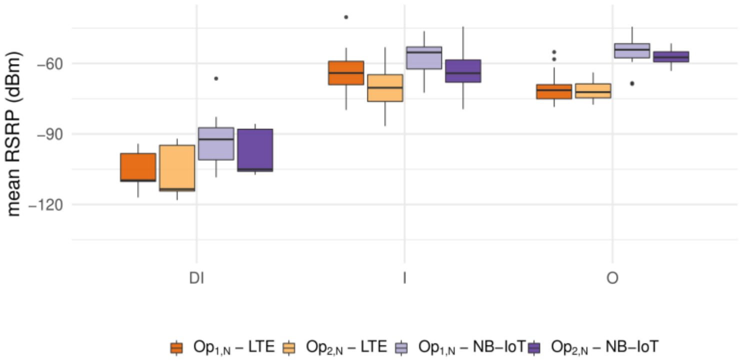
We conducted measurements in various areas of the city and different scenarios: Deep indoor (DI), for basements and deep enclosed spaces, Indoor (I), for houses and multi-floor buildings, and Outdoor (O), for outdoor while walking or on public transport. We further replicated a subset of our measurements over time (i.e., morning vs. afternoon vs. evening, and week vs. weekend), to account for temporal effects.

We considered the Reference Signal Receive Power (RSRP) [dBm] as a key metric for the NB-IoT RAN coverage KPI. We also collected Signal to Interference plus Noise Ratio (SINR [dB]), and Reference Signal Received Quality (RSRQ [dB]) as complementary measurements, along with the same three metrics for LTE. Finally, as reported in Annex Table A-46, Annex Table A-47 and Annex Table A-48 we defined a test case for the NB-IoT RAN Coverage KPI following the 5GENESIS experimentation methodology.

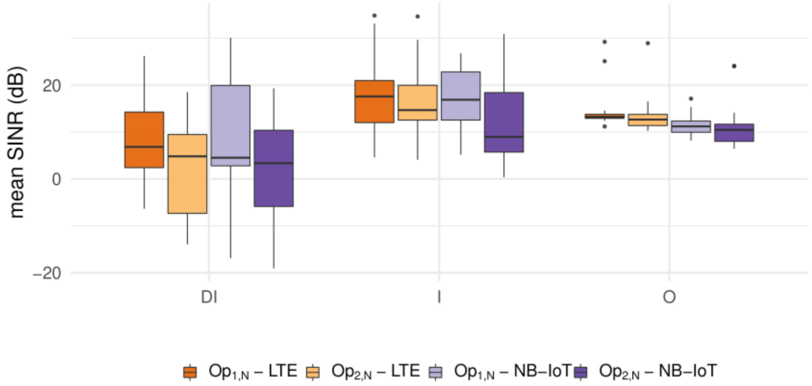
We then analysed NB-IoT RAN coverage for the two detected operators, showing how it changes across different scenarios and exploiting the LTE results for comparison. Figure 6-11 depicts the distribution of average RSRP, SINR, and RSRQ in a boxplot format with campaigns grouped per scenario (DI, I, and O) and coloured by operator and technology (NB-IoT and LTE).

Compared to LTE, NB-IoT provides significant RSRP boosts of 11.73, 7.87 and 15.01 dB on average for each scenario, respectively. This result is in line with the power boosting expected by 3GPP TS 36.104, which is of at least +6 dB when evaluated as the difference between the power of the entire NB-IoT carrier (180 kHz) and the average power over all carriers (LTE and NB-IoT).

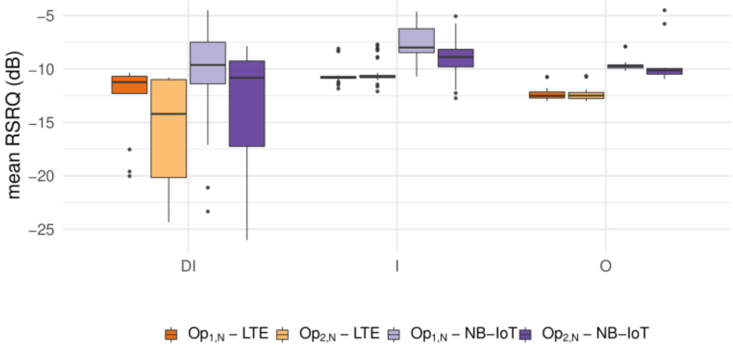
We also compare NB-IoT average RSRP across scenarios. In particular, the two operators show an average increase of 36.36 and 35.70 dB in terms of RSRP when comparing I to DI scenarios. This shows the negative effect of DI environments on signal propagation, which needs to be compensated by NB-IoT coverage enhancement techniques e.g., signal repetition. The deviation between O and I scenarios is instead reduced, with an average increase of 1.43 dB and 5.82 dB for the two operators, respectively.



(a)



(b)



(c)

Figure 6-11 Results of NB-IoT RAN Coverage test case: RSRP (a), SINR (b), and RSRQ (c).

7. BERLIN PLATFORM EXPERIMENTS

7.1. Overview

Phase 2 of the Berlin Platform experiments and trials aimed at conducting new performance evaluations addressing KPIs not covered in the previous phase, as well as running an initial field test during the Festival of Lights 2019.

In particular, the goals of the Berlin Platform for Phase 2 are to:

- gain experience with installing a nomadic, temporary 5G edge-deployment at Humboldt University during the Festival of Lights 2019,
- conduct an initial set of experiments, in particular focusing on evaluating video-related metrics, in a life-environment during the Festival of Lights 2019,
- extend the evaluation of 5G KPIs addressed by the Berlin Platform,
- extend the calibration tests started in Phase 1 to reflect the performance of newly installed compute and storage equipment,
- conduct interoperability tests of the Open5GCore with commercial-off-the-shelf (COTS) 5G SA equipment, and to
- execute initial 5G E2E performance tests.

Table 7-1 summarizes the components planned and available for the second integration cycle of 5GENESIS. All components planned for this cycle were available; minor upgrades scheduled for the upcoming phase of the project are required to address incompatibility issues between the OpenStack version deployed in the Berlin Platform and the Open5GENESIS Experimentation framework, which did not affect the experiments planned for Phase 2 of the project.

With the completion of Phase 2, the Berlin Platform finished the assessment of the following 5G KPIs:

- Service Creation time (as part of Phase 1, c.f. deliverable D6.1 Section 11.1.9), and
- User Density.

In addition, we conducted measurement to assess:

- Delay (Round-Trip-Time).
- Throughput.
- End-to-End GOOSE Message Latency.
- Evaluation of smart grid control traffic.

The former two involve initial assessments of an end-to-end 5G Stand Alone connection, thus exceeding the contractual obligation per DoW.

Table 7-1 Experimentation methodology components in the first integration cycle for the Berlin Platform, according to deliverable D2.3 [1]

Experimentation methodology component	Plan for integration and trial Phase 2	Status / Trial Phase 1 achievements
Open API's + Dispatcher	Yes	The Berlin Platform supports a set of test cases via the OpenAPI below the 5GENESIS Portal
Experiment Life Cycle	No	-
Portal	No	The Berlin Platform deployed the 5GENESIS Portal and executed tests via the portal
Custom experiments	POC	The Berlin Platform supports custom experiments over an E2E 5G SA network core. 5G SA COTS RAN (partially pre-production releases) was integrated
Standard experiments	Yes	Standard experiments may be executed via the Portal
E2E slices	No	-
VNF's	(Yes, Phase 1)	The Berlin Platform supports dynamic placement of VNFs in the testbed. Placement of VNFs may be triggered by the orchestration tool as well as by Keysight's TAP, which is chosen in 5GENESIS to execute and control experiments.
Scenarios	POC (mmWave backhauling)	The instantiation of the Berlin Platform for the first trial provides mmWave backhaul links (established in the lab as well as during the Festival of Lights 2019 field trial).
Un-attended experiments	POC	All experiments conducted in Phase 2 on the Berlin Platform are fully automated and controlled by TAP. The only exception were the video related measurements conducted during the Festival of Lights 2019 field trial.
Attended experiments	POC	The Berlin Platform supports in the lab *unattended* and attended experiments. Video-related experiments were executed as attended tests
Security Manager	NA	

Table 7-2 Primary 5G KPIs evaluated at the Berlin Platform in the first trial

KPI to be evaluated at the Berlin Platform according to DoW	Evaluated in Phase 2 / Second Trial	Comment
Density of Users	yes	Tests were done to evaluate the max. number of users in one cell / registered to the Open5GCore using commercial Keysight tools as well as the Open5G Benchmarking Tool
Service Creation Time	(Phase 1)	Evaluated in Phase 1 and reported in D6.1
Speed	no	Not scheduled for Phase 2
Reliability	no	Not scheduled for Phase 2
Additional 5G KPIs evaluated at the Berlin Platform	Evaluated in Phase 1 / First Trial	Comment
RTT	yes	-
Throughput	yes	-
E2E GOOSE Message Latency	yes	Special KPI to assess the video performance during the Festival of Lights 2019 field trial

7.1.1. System Architecture

All experiments were executed on the instantiation of the Berlin Platform as illustrated in Figure 9-1 of D6.1 [11]. The overall testbed was extended towards the main building of the Humboldt University (HU), which was connected by replicating the VPN-based interconnection approach between FOKUS and IHP thus extending the Berlin Platform to include three sites: Fraunhofer FOKUS, IHP, and HU.

Details of the new part of the system architecture, including the Festival of Lights deployment are described next as well as in Section 2.1 of D4.14 [10]. The architecture at FOKUS and IHP remains unchanged except the replacement of switches (new manufacturer) and installation of 5G SA RAN at FOKUS and replacing the 60 GHz link with upgraded revision of the 60 GHz technology at IHP.

7.1.2. Set-up for the Festival of Lights 2019 Field Trial

For the Festival of Lights 2019, the 5GENESIS Project deployed a nomadic, edge-based 5G core network at the courtyard of HU. The edge deployment was located in a car (van) that features a commercial, data-center-graded compute and storage infrastructure (NetApp HCI system), which was used to host the Open5GCore and the video processing software. The video camera, producing a 360-degree view of the illuminations at this festival, was connected to the

Open5GCore using the Open5GCore's UE/gNB emulation as 5G SA RAN equipment was not available during the field trials conducted in 2019. Test users could access the system via a deployed Wi-Fi access network. The edge was connected via a 60 GHz backhaul to the GÉANT network towards the main data-center at the FOKUS site, which allowed for accessing the video data processed in the edge from a (public) Internet access.

With the exception of using an actual 5G SA new radio system, the deployment of the field trial represented a full E2E network deployment, in which all data streams were forwarded via a full 5G SA core network.

The goal of this first trial was to gain experience with the nomadic, edge-based deployment at HU and to obtain first experiment results. The use of a full 5G SA RAN is planned for the final trial phase scheduled for 2021 (shifted from 2020 to the next year due to COVID-19 restrictions).

A full animation of the deployed system of the system set-up is available as a YouTube video ("Portable 5G Network: Field Trial at the Festival of Lights 2019 in Berlin", see https://www.youtube.com/watch?v=DBSs96DgWf0), which was also disseminated to nearly 200 participants from industry during the FUSECO Forum 2019. A snapshot of the video is shown in Figure 7-1, and the setup and the 5GENESIS team involved in the trial are shown in Figure 7-2.

Figure 7-1 Visitors of the Festival of Lights 2019 Field Trial with the Portable 5GENESIS Deployment

Figure 7-2 Berlin Platform 5GENESIS Team with the Portable 5GENESIS Deployment used during the Festival of Lights 2019

7.1.3. 60 GHz Backhaul Lab Set-Up

During Phase 2 of the project, IHP's 60 GHz units evolved to a new revision of the backhaul link that features beam steering. Although there was a plan to deploy a pair of nodes at the rooftop of IHP for this deliverable (c.f. Section 2.2.1.3 of D4.14 [10]), this was not possible due to the COVID-19 pandemic. Therefore, the millimeter wave (mmWave) 60 GHz wireless link was tested indoors, in a laboratory, at IHP premises. The deployment of the systems followed the system architecture described in Figure 9-1 in Section 9.1 of D6.1. Figure 7-3 shows the indoor set-up used to execute the performance measurements in Phase 2 of the project.

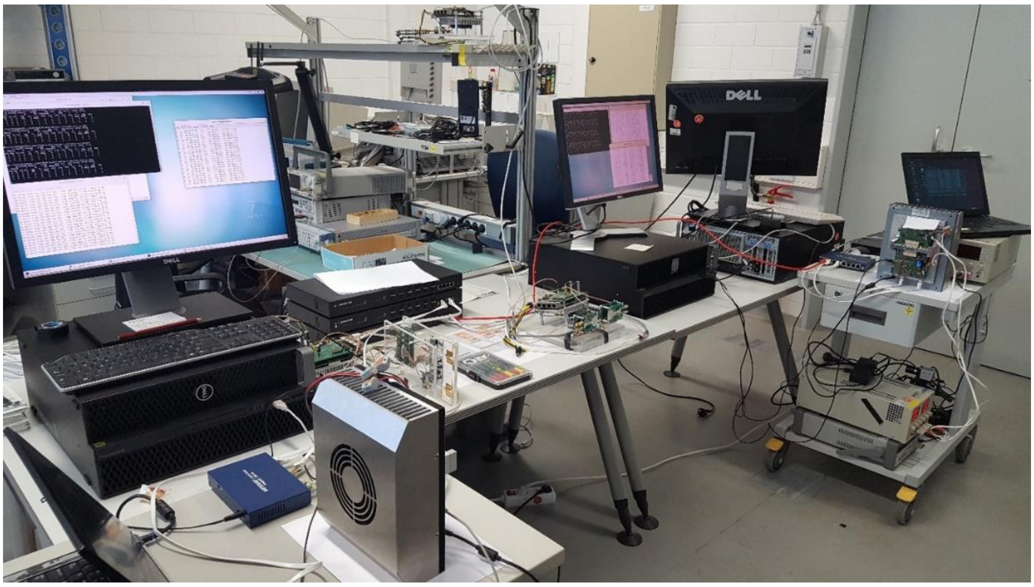


Figure 7-3 Testing of the mmWave wireless link consisting of two IHP's 60 GHz devices (2nd generation)

7.1.4. Summary of Key Features of the Berlin Platform

As initially scheduled for the overall duration of the Phase 2 of the project, the Berlin Platform provides an initial deployment of:

- A portable 5G (core), edge-based deployment of the testbed during the Festival of Lights 2019.
- An enhanced 60 GHz backhaul system featuring beam steering.
- Integration of (partially pre-commercial) 5G SA RAN in the testbed.

Experiments conducted during the first field trial during the Festival of Lights 2019 and in the Berlin lab facilities provided a performance assessment of

- A 360-degree video system running over a 5G SA network core.
- 5G SA new radio hardware of various manufacturers.

The experience gained from Phase 2 shows that the system concept for the portable, edge-based testbed extension of the Berlin Platform at HU is well capable of running field-trials in public environments. Moreover, initial results from the evaluation of the 5G SA RAN attached to the Open5GCore show that, with the final release of commercial 5G SA RANs, the Berlin Platform is capable of achieving the target KPI values for delay, throughput, and user density, as they were set as the target in the DoW.

7.2. Experiments and results

7.2.1. Festival of Lights 2019

7.2.1.1. Throughput

The throughput experiments aimed at evaluating the connectivity between the involved sites, i.e. Fraunhofer FOKUS and the HU, as well as to assess the observed throughput to the user end-devices.

The GÉANT-based connection between the two sites features an average achievable throughput of approx. 465 Mbps; the local backhaul, which connects the nomadic set-up in the courtyard of HU, allows for 830 Mbps. Thus the GÉANT connection is the limiting factor when connecting the data-center at FOKUS to the remote installation. Thus, **with respect to the experienced throughput, a data-center-based installation of the Open5GCore would have been feasible. However, for the experiments** during the Festival of Lights 2019, **an edge-based deployment was installed**. Regarding the observed **throughput imposed by the link towards the end-systems, approx. 40 Mbps were observed,** which is well **suitable for the video experiments conducted** during the event. Figure 7-4 shows the average throughput achieved over the different network segments during the Festival of Lights 2019.

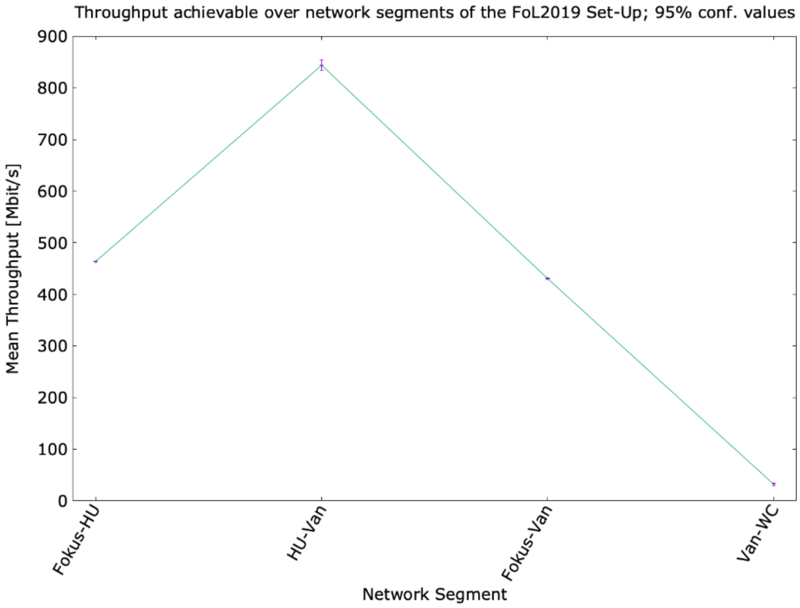


Figure 7-4 Throughput achieved over different network segments during the FoL-2019 field trial

Annex Table A-49 provides the detailed measurement results.

7.2.1.2. E2E RTT

In contrast to the throughput achievable between FOKUS and the remote installation at the HU, the experienced RTT mandates an edge-based installation for low-latency applications. While the GÉANT-connection imposes on average a 3 ms RTT, the backhaul system used for connecting the edge-installation at the courtyard has an average RTT of 10 ms. Noticeable, the variations in the RTT go even up to 200 ms at most, which makes a purely edge-based deployment favourable for low-latency use cases. Figure 7-5 depicts the E2E delay (RTT) achieved over different network segments during the Festival of Lights 2019 field trial.

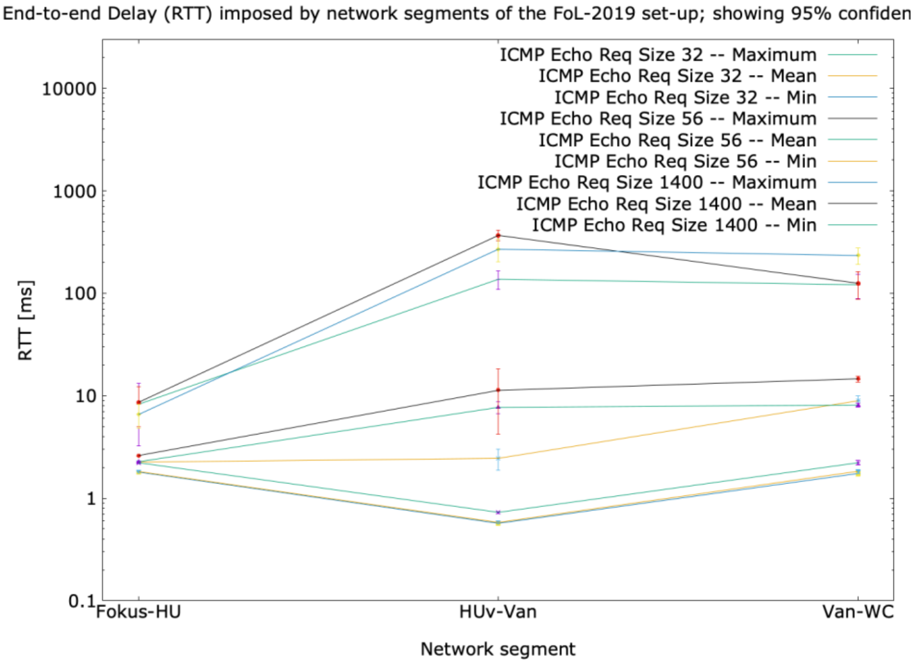
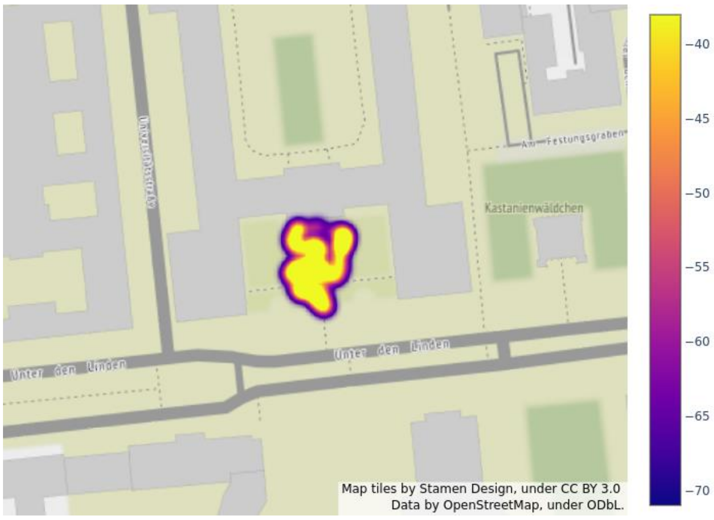


Figure 7-5 RTT achieved over different network segments during the FoL-2019 field trial

Annex Table A-50 provides the detailed measurement results.

7.2.1.3. RAN Coverage

The access network provided to the end-users during the experiments was rather small in size. Though the actual coverage of the network was larger, location reports used to sketch the RAN coverage in Figure 7-6 were received from users located within the courtyard (approx. 500 m^2) of HU. As such, for the final experiments to be conducted in Phase 3 of the project, assessing the RAN coverage should not rely solely on test users, but should include manual measurements of team members exploring the maximum reach of the user connectivity provided by the nomadic installation.



Map tiles by Stamen Design, under CC BY 3.0
Data by OpenStreetMap, under ODbL.

Figure 7-6 Festival-of-Lights 2019 network coverage

7.2.1.4. 360° Video Streaming

For the 360° Video Streaming experiments, four different tools were used to collect the data. Each one dealt with its own kind of data and its own part of the experimental setup (i.e. client, server, or network):

- A custom-built JavaScript/PHP web application that captures data on web sessions, geolocation, and network connection on (primarily) the client side.
- Bitmovin Analytics -- an API-driven video analytics system that provides insight on player performance, user behaviour and other aspects of the entire video chain.
- Wireshark, which is a free and open-source packet analyzer. The application was configured to run as a background process on the web server, capturing several network and transport layer packet fields from incoming traffic.
- OpenTAP -- a free and open-source test sequencer -- used in conjunction with custom test scripts to perform throughput and latency measurements on different lags of the network.

Five different criteria for quality assessment are considered, the analysis of which primarily relies on data collected by Bitmovin Analytics. Data collected by other means have been used as corroboratory material where appropriate:

- Average times spent on different player states.

- Total times spent playing on different bit rate levels.
- Rebuffering event frequency.
- Start-up times per impression.
- Quality switch frequency.

Two sets of data were collected during the Festival of Lights 2019:

- *BERLIN I:* The first round of experiments was conducted between 2.00 p.m. and 10.00 p.m. on Friday, October 18, 2019. A small group of testers were recruited among the people on the site, who were either directly involved with the project or otherwise informed about its purpose. The testers were instructed to use their smartphones or tablets and connect to a private Wi-Fi hotspot, which provided them with access to the internet. Using their web browser of choice, the testers then navigated to a URL that pointed them to the local web server that displayed the 360° live video stream. Unfortunately, an issue related to the video frame size caused suboptimal results at this stage.
- *BERLIN II:* The second round of experiments was initiated on Saturday, October 19, 2019 at 6.00 p.m., and concluded the following morning, on Sunday, October 20, 2019 at 10.00 a.m. The video frame size issue (and a couple of other less critical bugs) had already been resolved, and the system was operating in a satisfactory manner -- at least as far as the eye could tell. The decision was therefore made to allow open access to the 360° live video stream, which was provided through a public Wi-Fi hotspot. Two circumstances somewhat limited the amount of collectible data: (i) printed advertisements and other public announcements were avoided for privacy concerns, and (ii) for about the last half of the period, outbound internet access was blocked off from within the walled garden of the Wi-Fi hotspot, and hence, the cloud-hosted analytics engine was unreachable.

Table 7-3 shows some key figures from the two rounds of the experiments. These figures measure the size of available data. By all measures, more data was collected on the first day. As mentioned above, however, these data were tainted by an issue related to the video frame size, which caused the player to behave poorly on certain platforms. Such this, data from the first round of experiments are perhaps less interesting on their own than that from the second round. It is also worth mentioning that clients were manually assigned static IP addresses on Friday, as these were testers involved with the project, while IP addresses were handed out by DHCP to the public during the second round. Only the first IP address in the DHCP pool was ever utilized (i.e. at most one client was connected to the Wi-Fi hotspot at any one time).

Table 7-3 Key Figures from Video Experiments

	Berlin I (Friday)	Berlin II (Saturday)
Unique impression IDs	105	32
Unique user IDs	28	13
Unique IP addresses	7	1
Unique user agents	11	6
Total play time (in seconds)	5862	3310

Next, we analyze the results for both datasets. However, we focus the discussion on the results of the Berlin II dataset, as Berlin II data is collected when the video streaming system was working as intended and expected.

In Figure 7-7 for Berlin I, we observe that, on average, the player spent 0.2 seconds on setup, 131.9 seconds on start-up, 3.9 seconds on rebuffering, and 12.4 seconds on playing. No measurable time was spent on seeking, quality changes, or errors. On the other hand, for Berlin II, on average, the player spent 0.2 seconds on setup, 2.2 seconds on start-up, 0.7 seconds on rebuffering, and 12.6 seconds on playing. No measurable time was spent on seeking, quality changes, or errors.

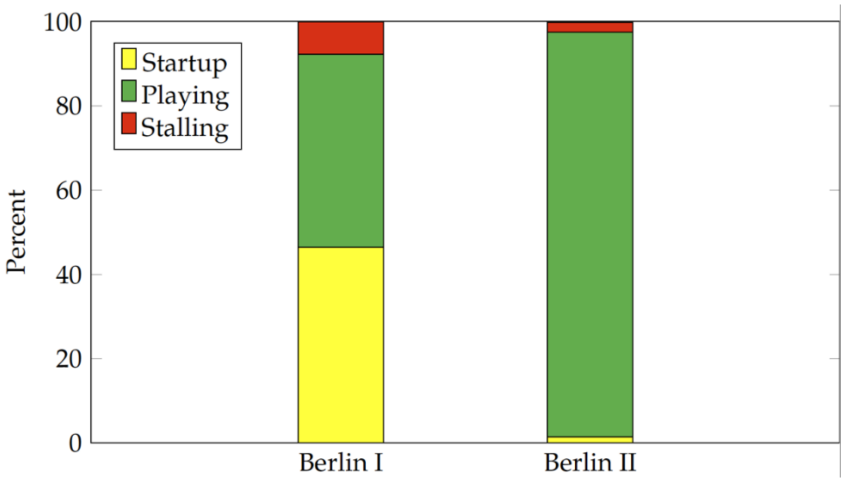


Figure 7-7 Average Times Spent in Different Player States

The average time spent in different player states show that across all impressions, the player was indeed playing the live video stream more than 75 percent of the time for Berlin II. The average total duration of one single impression equates to about 16 seconds. Note that this holds as long as pauses are discounted, which occur, for example, on a smartphone when the web browser is put in the background or when the screen is turned off. Comparatively longer average time is spent on pauses (quite naturally), but this observation is particularly interesting. The average setup and start-up times were quite short, less than one second and slightly more than two seconds, respectively, and in the course of one entire impression, the player spent just an average of just below one second rebuffering. No measurable time was spent on seeking, quality changes or errors, and these states are also not shown in Figure 7-7. In Figure 7-8 we illustrate the total times spent playing different bit rate levels. For Berlin I, in total, the player spent 446.6 seconds on the 1000 kbps level, 2389.9 seconds on the 4000 kbps level, 979.9 seconds on the 6000 kbps level, and 1080.2 seconds on the 8000 kbps level. For Berlin II we have adjusted the quality levels as illustrated in the figure. We observe that, in total, the player spent 70.9 seconds on the 1100 kbps level, 23.4 seconds on the 4400 kbps level, and 765.5 seconds on the 6600 kbps level.

Higher bit rate provides higher video quality, and thus higher Quality of Experience (QoE). The total times spent playing different bit rate levels show a strong preference for the highest level. This is no surprise since only one client was watching the video stream at any one time for the duration of the whole round. The network was never congested, and the bandwidth was more than sufficient to supply the highest quality. Comparatively shorter times were spent on the middle and lowest bit rate levels, which is likely an effect of the adaption algorithm on the client side.

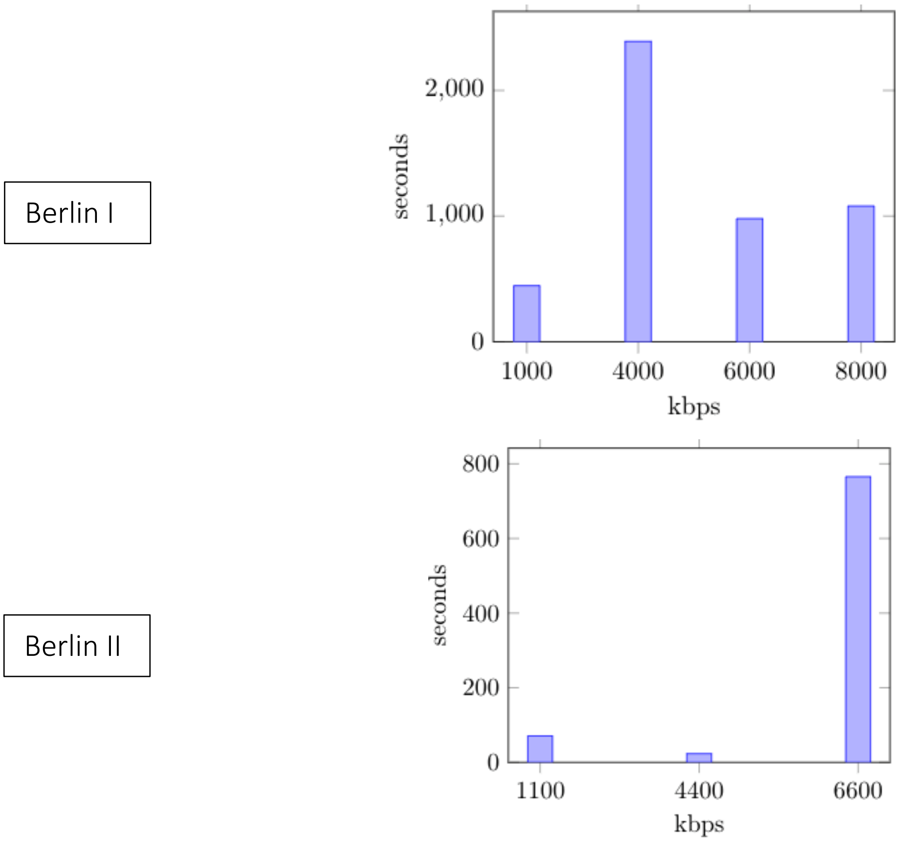


Figure 7-8 Total Times Spent Playing Different Bit Rate Levels

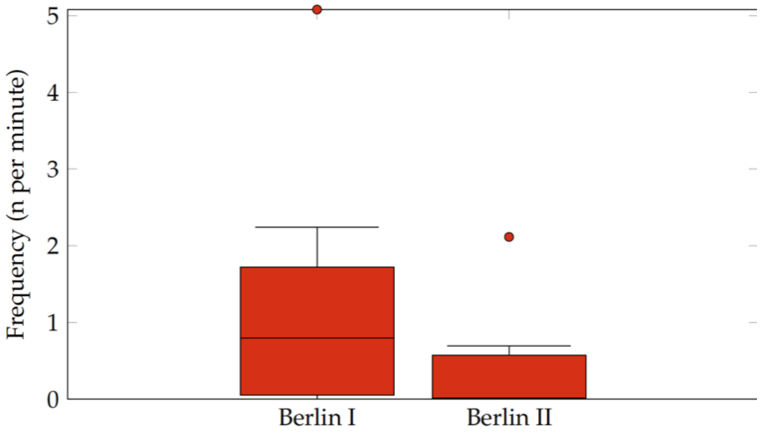


Figure 7-9 Rebuffering Events Frequency

In Figure 7-9 we present the rebuffering event frequency (number of stalls per minute). For Berlin I, we observe that the stall per minute were distributed around a mean of 0.75 while for Berlin II, the mean stalls were almost 0.

Any one rebuffering event is undesirable, as it makes the player stall -- if only for a short while. A few rebuffering events at the beginning of the stream would most likely be tolerable. Many rebuffering events dispersed throughout the impression, however, drastically reduces QoE. For Berlin II, half of all impressions contained at most three rebuffering events, while 75~\% of them contained at most nine. Considering that each impression lasted various amounts of time, this is considered satisfactory. A few impressions contained a higher number of rebuffering events, which appears to be a function of mostly duration.

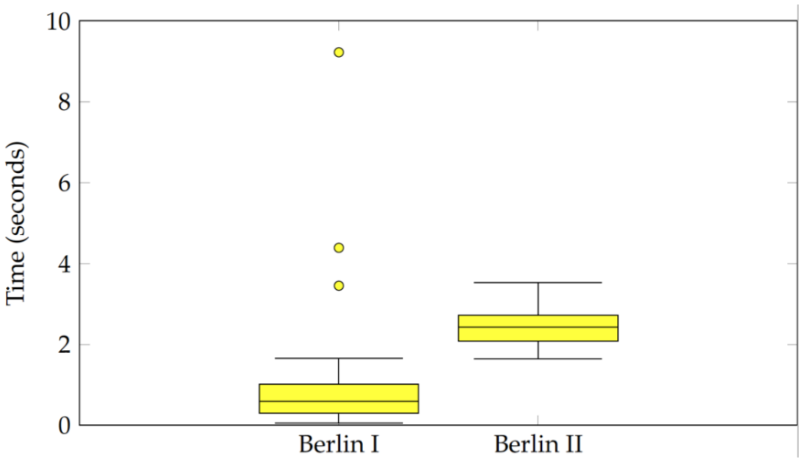


Figure 7-10 Start-up Times per Impression

In Figure 7-10 we illustrate the start-up times per impression. For Berlin I, the start-up times per impression (in seconds) were distributed around a mean of 0.6, with 1.2 and 0.3 as the third and first quartiles, respectively. The upper and lower whiskers were calculated above and below these limits at a distance of one and a half times the interquartile range, equaling 0.05 and 1.7, respectively. Three outliers are seen at 3.5, 4.4, and 9.2. Additional outliers not shown (too far off) are also found at 25.2, 33.6, 266.9, 5577.2.

For Berlin II, the start-up times per impression were distributed around a mean of 2.4, with 2.7 and 2.1 as the third and first quartiles, respectively. The upper and lower whiskers were calculated above and below these limits at a distance of one and a half times the interquartile range, equalling 3.5 and 2.0, respectively. No outliers are seen.

As short as possible start-up time is desirable, as it lowers the delay from when the client opens the web page to when the video stream is watchable; it provides a more responsive QoE. A delay of some seconds should be expected and accepted -- as this is a one-time cost in terms of QoE. For Berlin II, no impression took longer than 3.5 seconds to get started, which is well below any reasonable threshold for impatience. Start-up times across impressions were also very consistent, as the difference between the longest and the shortest delays measured was only 1.5 seconds. Also, there were no outliers outside either of these limits.

In Figure 7-11, we illustrate the Quality switch frequency per minute. For Berlin I, we observe that the mean quality switch per minute is around 0, while for Berlin II the mean quality switch frequency is 0.08.

Though not as bad as rebuffering events, quality degradations, i.e. switching from a higher to a lower bit rate level, adversely impacts QoE. Here, quality switches in the reverse direction are also counted. A few quality switches (typically at the beginning of the stream) are expected, as the player will often start carefully, and then gradually go braver if network conditions so permit. Sometimes, the player needs to fall back to a lower level again. It may take some trial and error before it arrives at a (temporarily) stable choice. This behaviour is apparent from the observation that most impressions contain a number of quality switches higher than or equal to the number of available bit rate levels. For Berlin II, the highest number of quality switches among all impressions, however, is merely five, so the behaviour is well contained. It is not erratic, as was the case in the first round of experiments.

Annex Table A-51 provides the detailed measurement results.

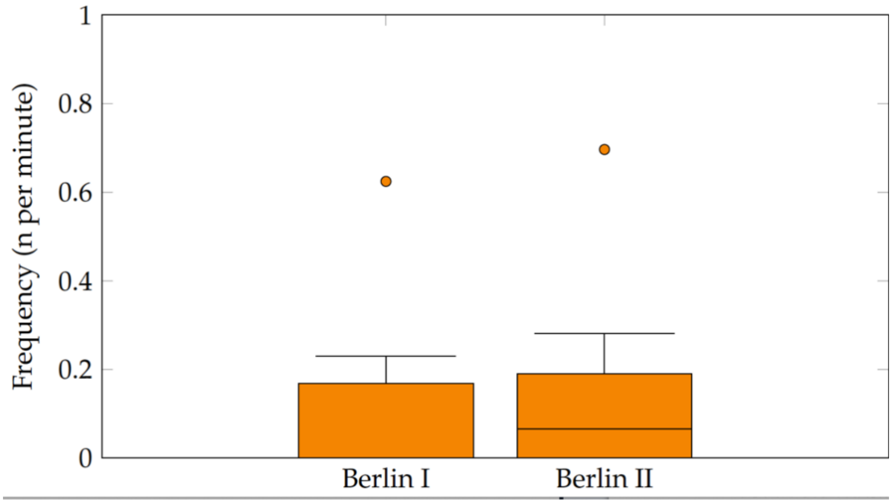


Figure 7-11 Quality Switch Frequency

7.2.2. OpenStack and NetApp HCI

The goal of these lab experiments was to gauge the performance limitations of the new, underlying hardware infrastructure at the 5G Lab at FOKUS. Such evaluation of the baseline infrastructure is of essential importance in order:

- to assure that further assessments of 5G SA RAN as well as the Open5GCore are truly exposing the performance limitations of the 5G system component and not the limitations of the underlying hardware, and
- to fully understand the performance limitations of the assessed 5G Core and RAN attached to it by observing the overall end-to-end system performance while considering the performance of the underlying hardware.

Results show that the underlying hardware does not impose any limiting factors with respect to RTT or imposed delays between systems. In particular, the delay in terms of RTT added by the system is insignificantly small – i.e., in the order of 1 ms for end-to-end-experiments – which allows to state that the measurements of the KPIs for a 5G SA end-to-end system are dominated by the 5G system itself and not the underlying network and compute & storage infrastructure.

7.2.2.1. E2E RTT

The deployed compute and storage system included in the OpenStack testbed set-up showed a consistently low RTT between components deployed on the system. Maximum round-trip-times always were below 0.5 ms, even going down to an expected mean of 0.2 ms. Except for occasional outliers, the experienced RTT is invariant against different packet sizes.

Figure 7-12 shows the round-trip-time between VNFs/VMs deployed within the NetApp compute platform at different compute nodes.

The round-trip-time slightly increases to 1 ms when a second external component is used for the system evaluation, i.e. when an intermediate switch is included in the communication path. Figure 7-13 shows the round-trip-time across the central DellSwitch.

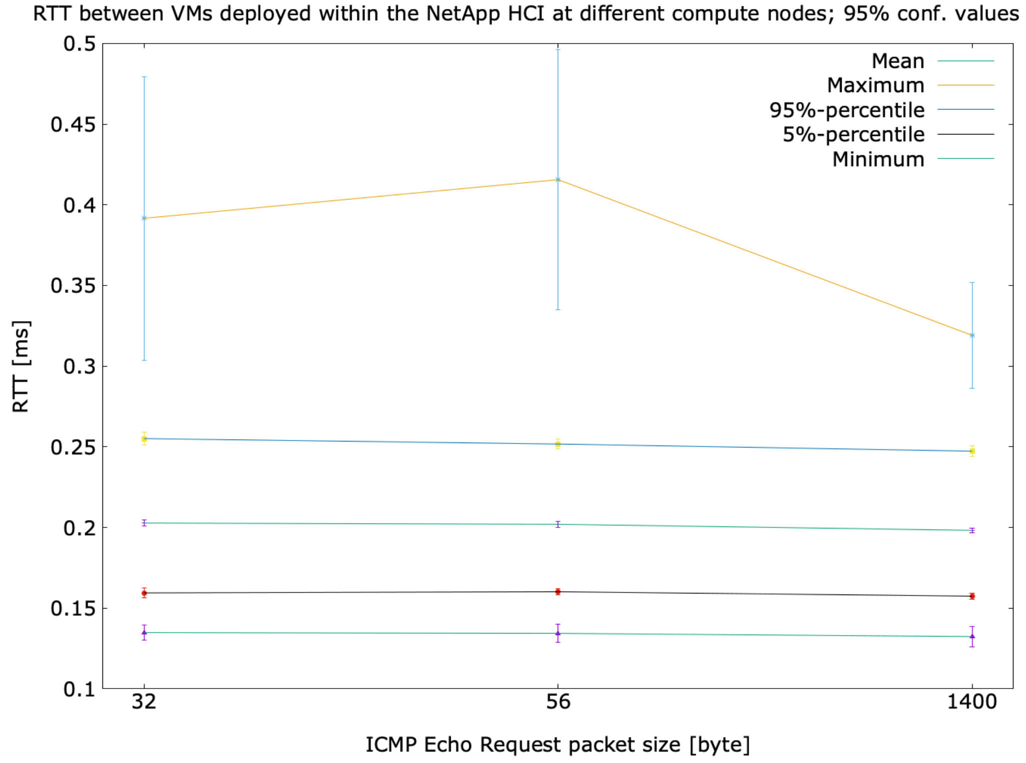


Figure 7-12 RTT between intra-storage-and-compute nodes

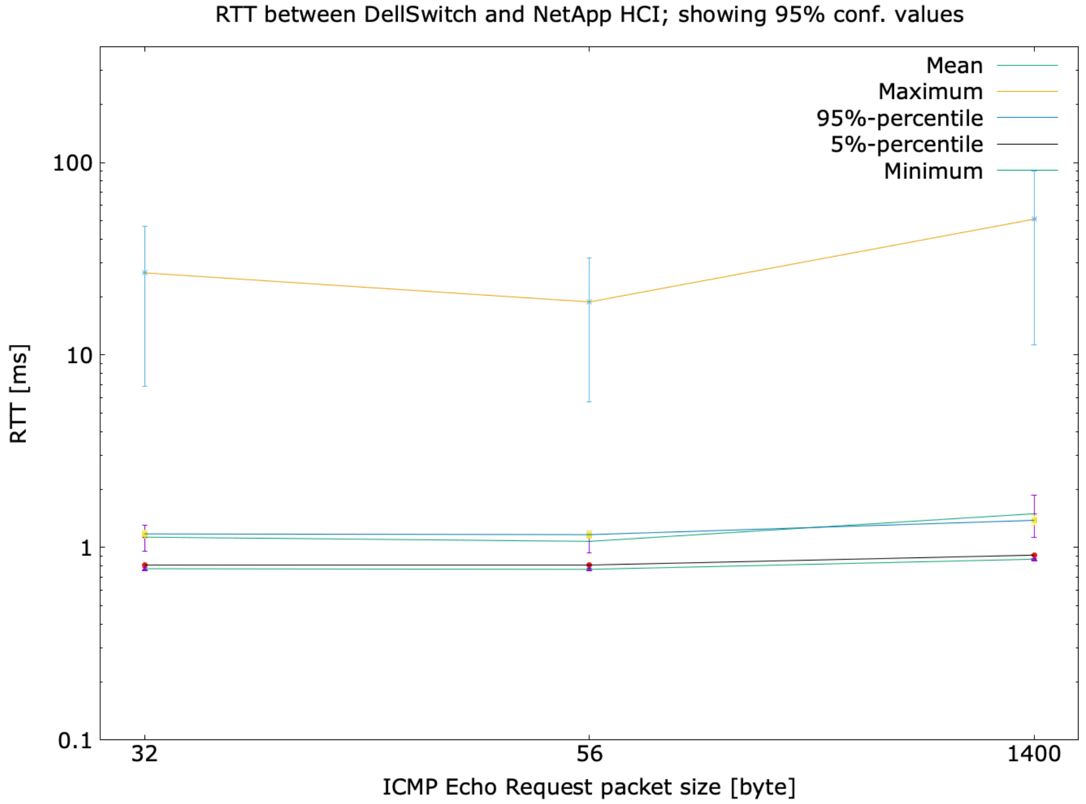


Figure 7-13 RTT across the central DellSwitch (between NetApp HCI and ThinkCenter)

Annex Table A-52 and Annex Table A-53 provide the detailed results for the tests.

7.2.2.2. Throughput

To assess the achievable throughput between components running on the same compute and storage systems, to groups of experiments were run. In one group, both, source and destination of the throughput test were located on the same compute unit (C0) within the system; whereas in the second group, tests were run between two compute units within the compute and storage cluster. Within each group, two experiments were run in which one (1) or four (4) parallel data streams were established between source and destination.

Figure 7-14 shows the accumulative throughput over all data streams within each experiment. Results show that achievable throughput when source and destination are on the same compute unit is slightly higher (19 Gbps) as when they are located on different compute units (14 Gbps). The accumulative throughput over four streams is even higher (28 Gbps vs 16 Gbps respectively).

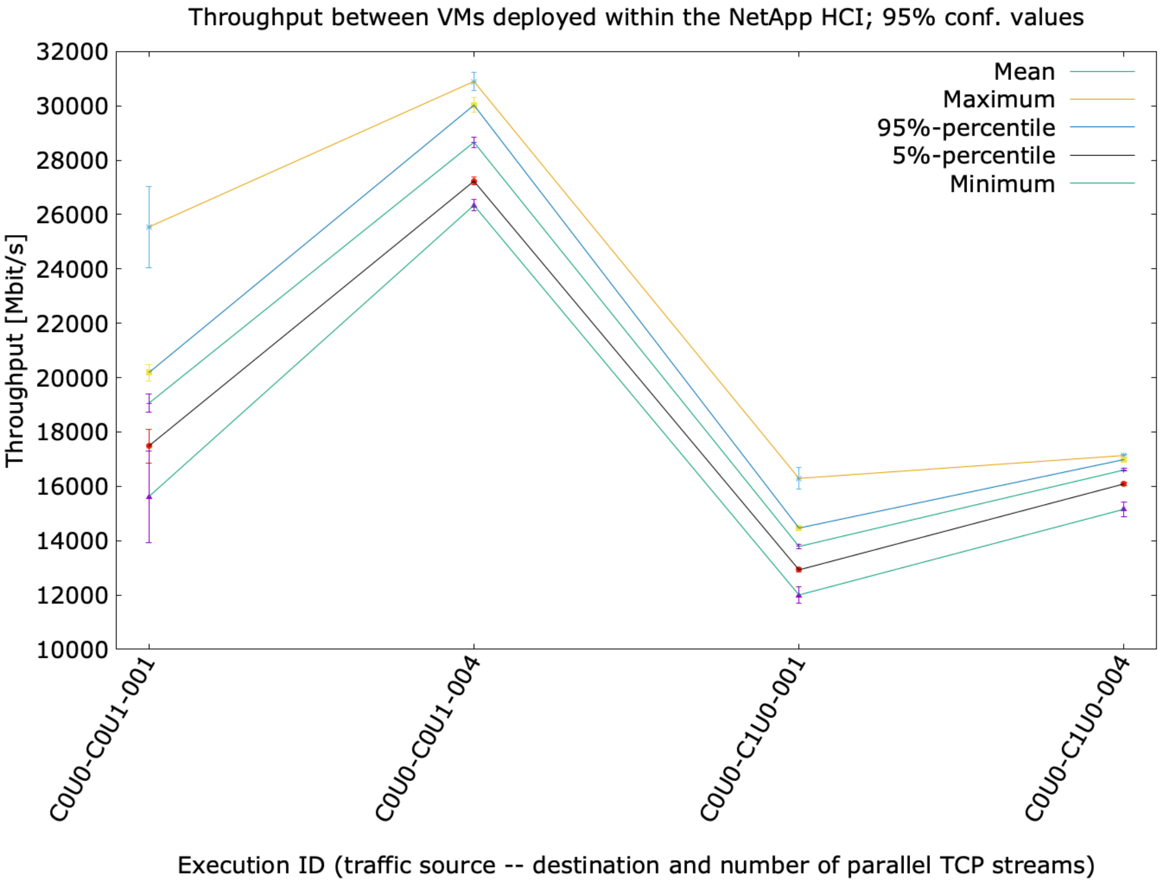


Figure 7-14 Throughput between intra-storage-and-compute nodes

For a second set of experiments, the achievable throughput between an external compute unit and the previously tested compute and storage unit is assessed (see Figure 7-15). The external compute unit, having 1 Gbps network interface card, is connected over a switch to the compute and storage system. The throughput for the up- and download direction is measured for having one and four parallel data streams. For all experiments, the throughput is symmetrical for the up- and downlink and reaches approximately 930 Mbps, which is close to the achievable network speed.

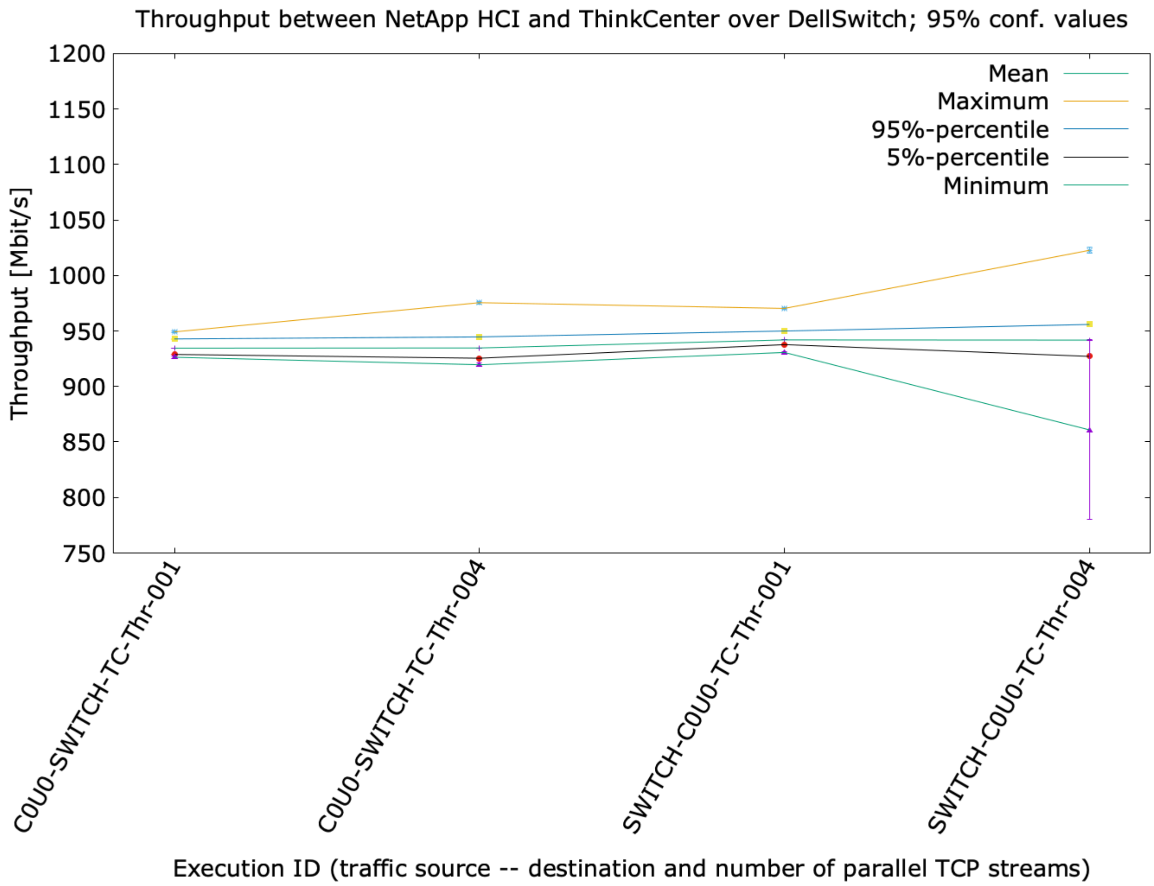


Figure 7-15 Throughput achieved over the central DellSwitch (between NetApp HCI and ThinkCenter)

The following tables provide the detailed test results for each test case: Annex Table A-54, Annex Table A-55, Annex Table A-56 and Annex Table A-57.

7.2.3. 5G Packet Core

The following evaluation of the Open5GCore use a virtual (emulated) gNB and UE, which is part of the Open5GCore distribution, or the commercial Ixia IxLoadCore testing tool. In all cases, results how that the Open5GCore SA Implementation is well suitable to handle low-latency, high capacity throughput of thousands of users, typical for small campus network deployments.

7.2.3.1. E2E RTT

Figure 7-16 shows the achievable RTT between an Ubuntu VM connected via a UE/gNB emulator (acting as a router) over the Open5GCore to an external server (Ubuntu VM). All components of the system under test run on the previously assessed compute and storage unit. The experienced RTT has a mean of 0.53 ms, which is invariant of different packet sizes, and is always below 0.8 ms (maximum). Thus, the Open5GCore is well suitable for providing ultra-low-delay 5G links.

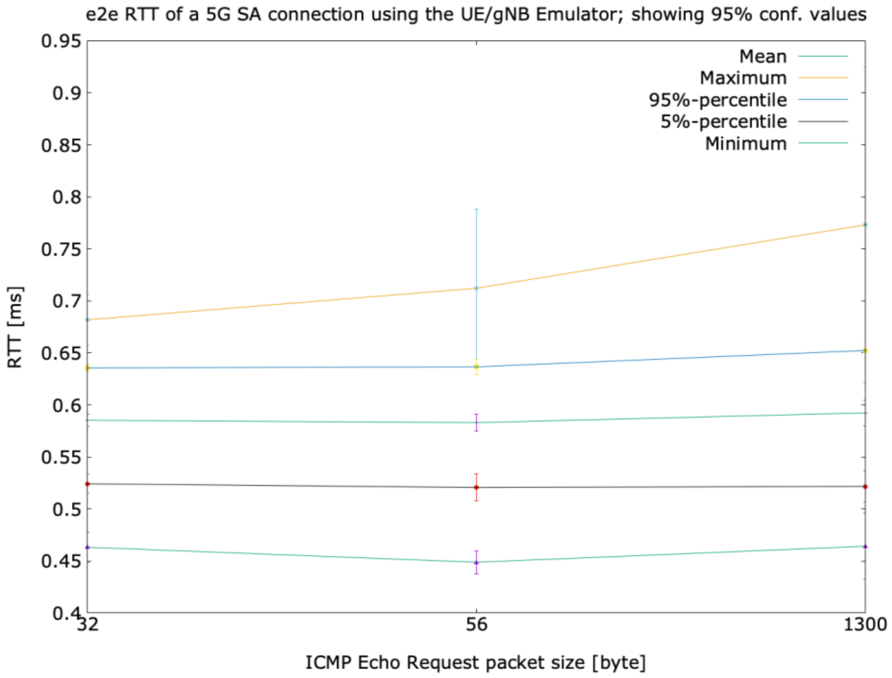


Figure 7-16 5G-SA end-to-end RTT (using the Open5GCore UE/NB Emulator)

The following table provides the detailed measurement results: Annex Table A-58.

7.2.3.2. Throughput

The same set-up previously used to assess the 5G end-to-end RTT is used to assess the achievable throughput. The mean throughput is in the order of 750 Mbps, reaching peaks of 880 Mbps and being always above 500 Mbps (see Figure 7-17). The core is well capable of handling several parallel user data streams which only insignificantly effect the overall accumulated system throughput.

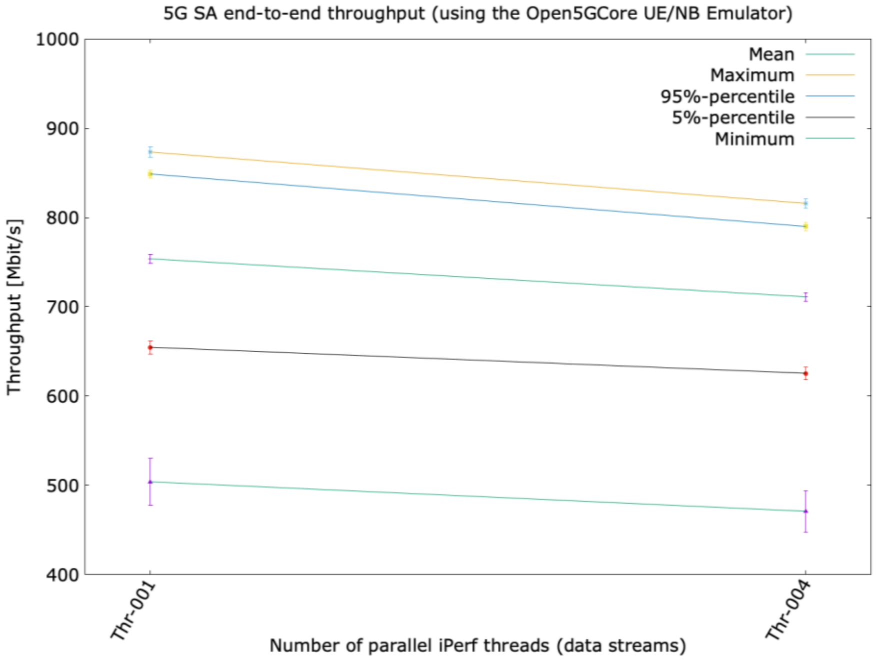


Figure 7-17 5G-SA end-to-end throughput (using the Open5GCore UE/NB Emulator)

The following table shows the detailed test results: Annex Table A-59.

7.2.3.3. User Density

Final tests evaluate the number of users that the Open5GCore can serve simultaneously. It should be noted that this indirectly translates to the supported user density, which depends apart from the number of supported users on the RAN's coverage area. Figure 7-18 shows the number of consecutively supported users attached to the system, before the registration of an additional user times out or produces an error. The tests are independently conducted using the Open5GCore Benchmarking Tool, as well as the commercial Ixia IxLoadCore test suit.

Both independent tests show that the Open5GCore can successfully handle approximately 2800 simultaneous UEs (a graphical view is shown in Figure 7-18). Considering that the user density KPI aims at supporting 10.000 devices per km^2 (1 device per 100 m^2), we can deduce that the Open5Gcore is suitable for handling a RAN coverage of 280.000 m^2, i.e. covering an area of approximately 530x530 meter, which is well suitable for small to medium campus network deployments, e.g. in a factory shop-floor. It should be noted that higher number of users can likely be supported by upscaling the compute and storage resources on which the Core is run; which is subject to further investigations in Phase 3 of the project.

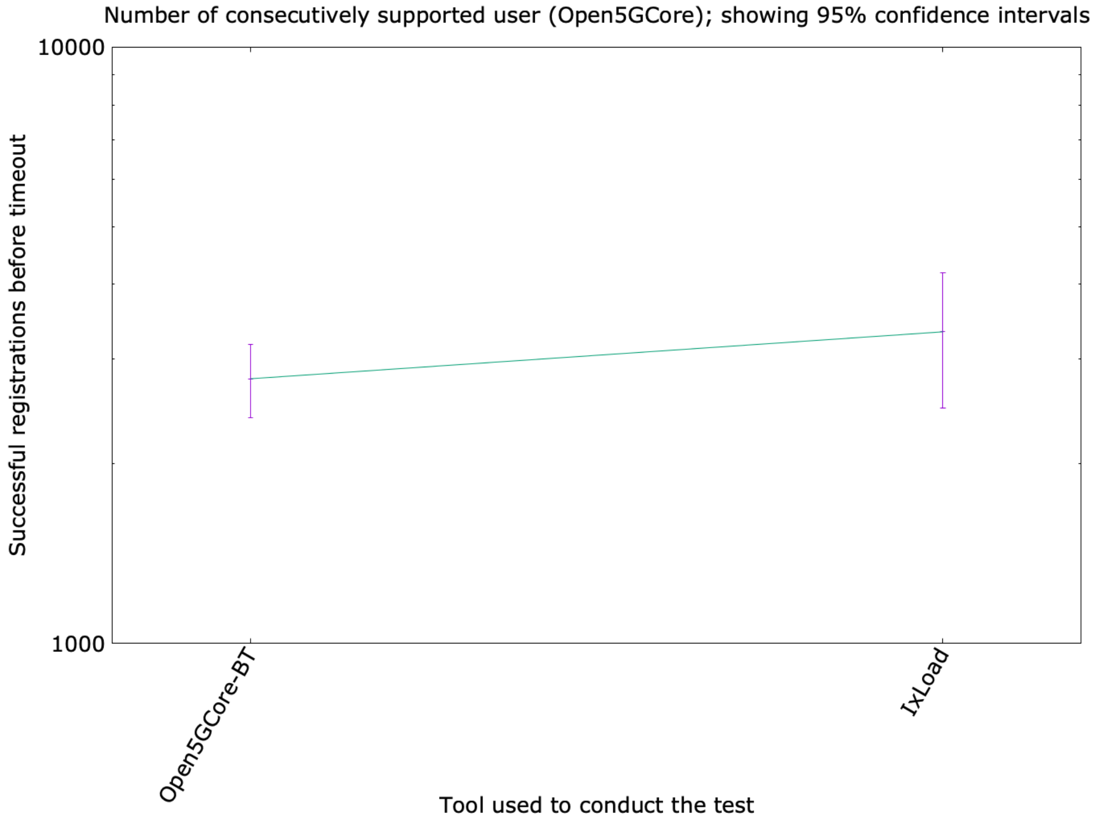


Figure 7-18 User Density: Number of consecutively supported user (Open5GCore)

The following table shows the detailed measurement results: Annex Table A-60.

7.2.4. Evaluation of 5G SA Cell Prototypes (multiple vendors)

As part of the 5GENESIS experiments, three different vendors of 5G SA equipment were evaluated regarding their E2E performance when attached to the Fraunhofer FOKUS Open5GCore. For all vendors, the evaluated equipment used pre-commercial / experimental firmware enabling 5G SA operation. Partially, Fraunhofer FOKUS was provided during the integration with daily builds of the firmware.

As pre-commercial equipment was partially used, results may not be reported in a public deliverable but are instead contained in a separate, confidential annex provided to the reviewers and the European Commission.

7.2.5. 60 GHz MetroLinq

The following results gauge the lab-based performance of the MetroLinq backhaul that was also used for the Festival of Lights 2019 field trial. While the lab-based results show ideal performance, well suitable for connecting field deployments of a 5G RAN to a core network deployed in a data-center, the previously described results from the field trial reveal significantly higher variations with respect to throughput and RTT, likely caused so sub-optimal placement of the antenna systems in a mobile, nomadic field trial. This underlines the need for edge-based architectures for nomadic trials.

7.2.5.1. RTT

As shown in Figure 7-19, the experienced RTT is always (in the lab) in the order of 1 ms, showing occasional outliers in the order of 40 ms.

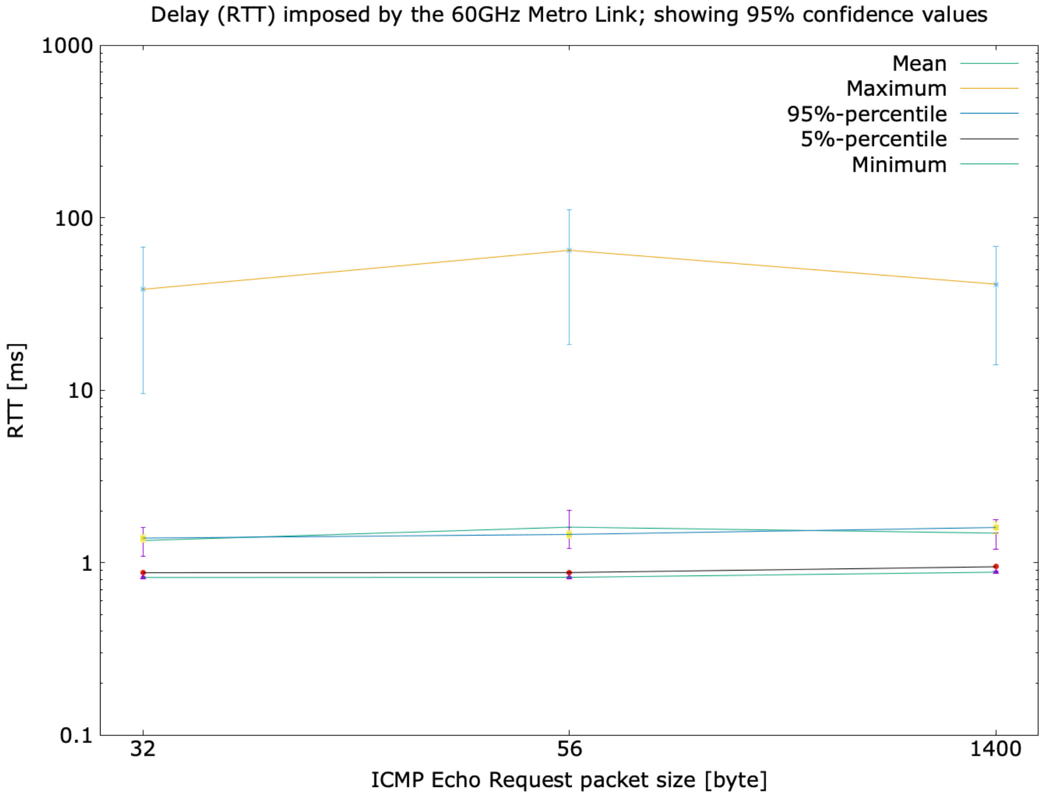


Figure 7-19 RTT between nodes interconnected via 60 GHz Metrolinq System

The following table shows the details measurement results: Annex Table A-61.

7.2.5.2. Throughput

The mean throughput of the system is always in the order of 930 Mbps; the system is invariant against the number of parallel data streams imposed on the system.

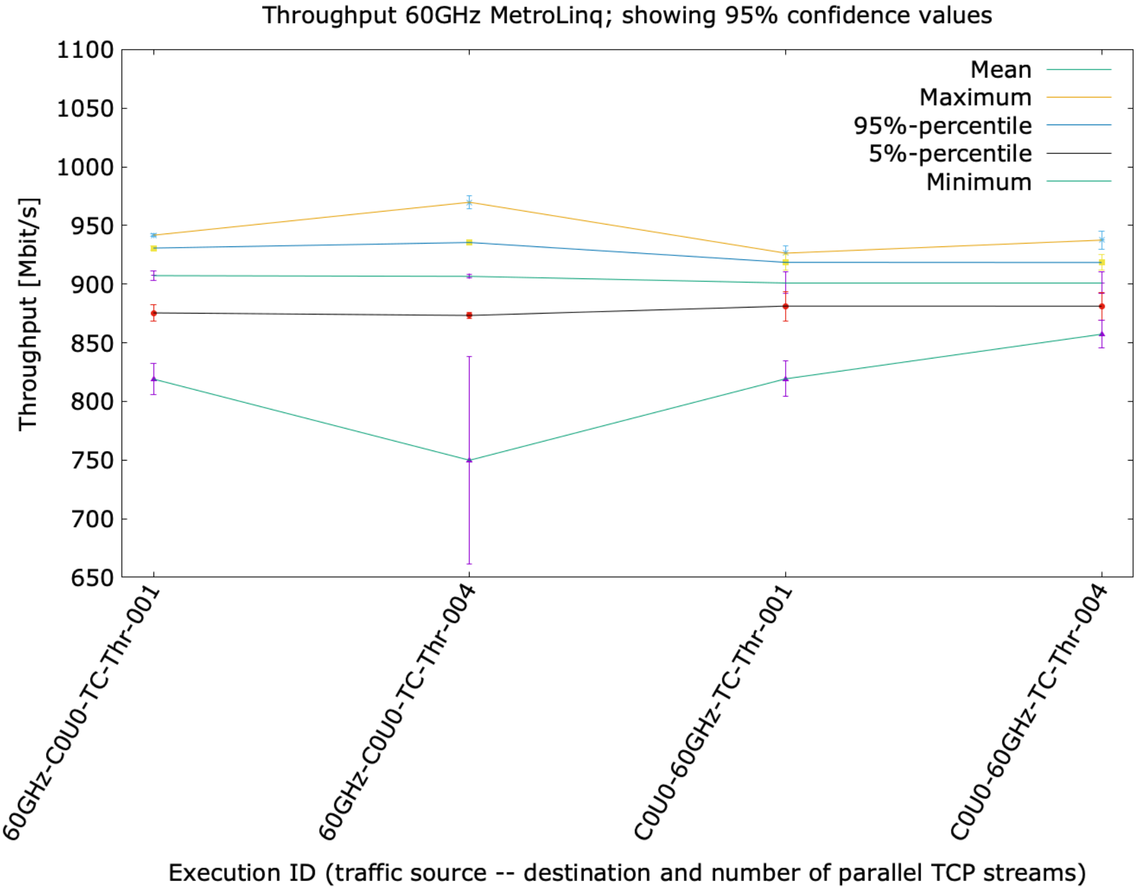


Figure 7-20 Throughput between nodes interconnected via 60GHz Metrolinq System

The following table provides the detailed measurement results: Annex Table A-62.

7.2.6. 60GHz IHP Prototype

7.2.6.1. RTT

The evaluation of the 60 GHz IHP backhaul system show a mean RTT of 0.4 ms for small and medium ICMP Echo Request packet sizes of 32 and 56 bytes. The experienced RTT slightly increases large, 1024-byte packets to approx. 0.6 ms. In all cases, the maximum experienced RTT was always below 0.9 ms (see Figure 7-21). For all measurements, the 95%-confidence intervals were below 0.05 ms. As such, one can conclude that the one-way delay imposed by the backhaul system is always well below 0.5 ms, which makes it a well-suitable backhaul technology for 5G edge deployments.

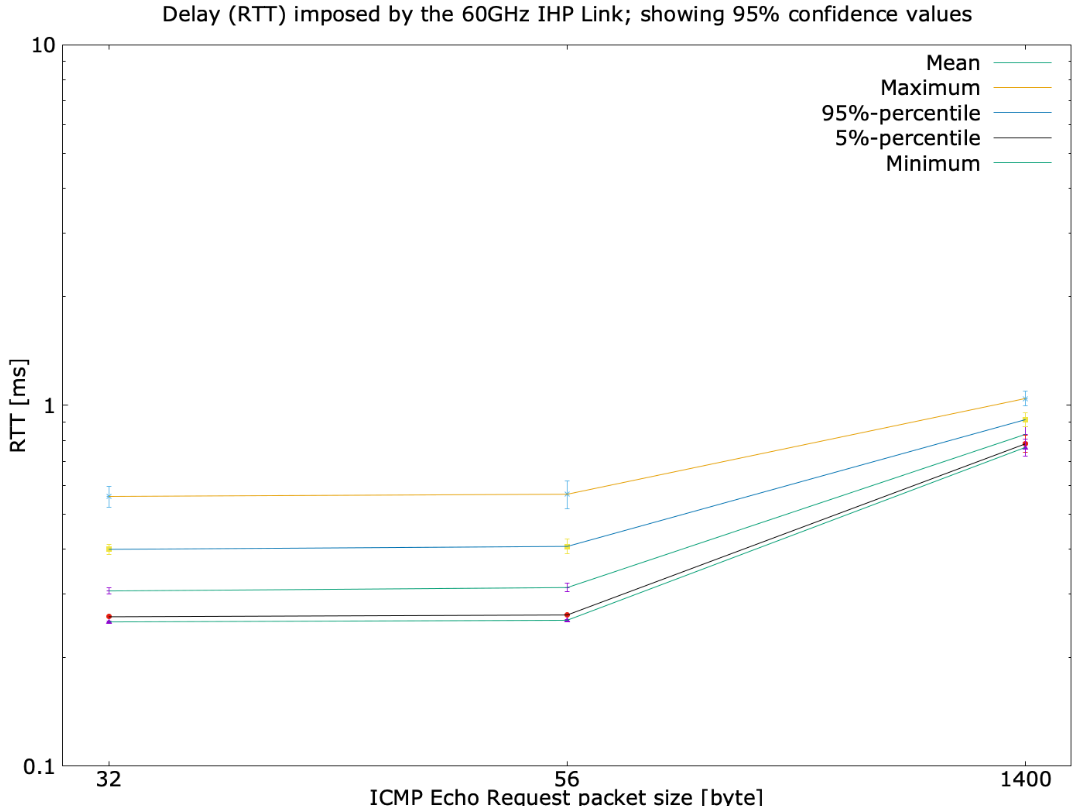


Figure 7-21 RTT between nodes interconnected via 60GHz IHP System

The following table shows the details measurement results: Annex Table A-63.

7.2.6.2. Throughput

The achievable throughput of the 60 GHz IHP backhaul systems varies for the uplink and downlink direction. In the uplink direction, the system features approx. 865 Mbps total throughput and for the downlink 890 Mbps (see Figure 7-22). The experienced throughput is nearly invariant against the number of parallel data streams and distributes the throughput evenly among the traffic streams. This shows that the implemented scheduling scheme of the system is fair. Results show that the system is suitable to up to two (2) 5G RANs to a 5G Core located remotely in a data-center (assuming 450 Mbps throughput achievable per RAN). At the same time, the experiments show that for high-capacity deployments involving several 5G RANs, a local, edge-based deployment of the network core is essential in order to support local processing of high capacity data, i.e. an edge-based deployment of data processing applications.

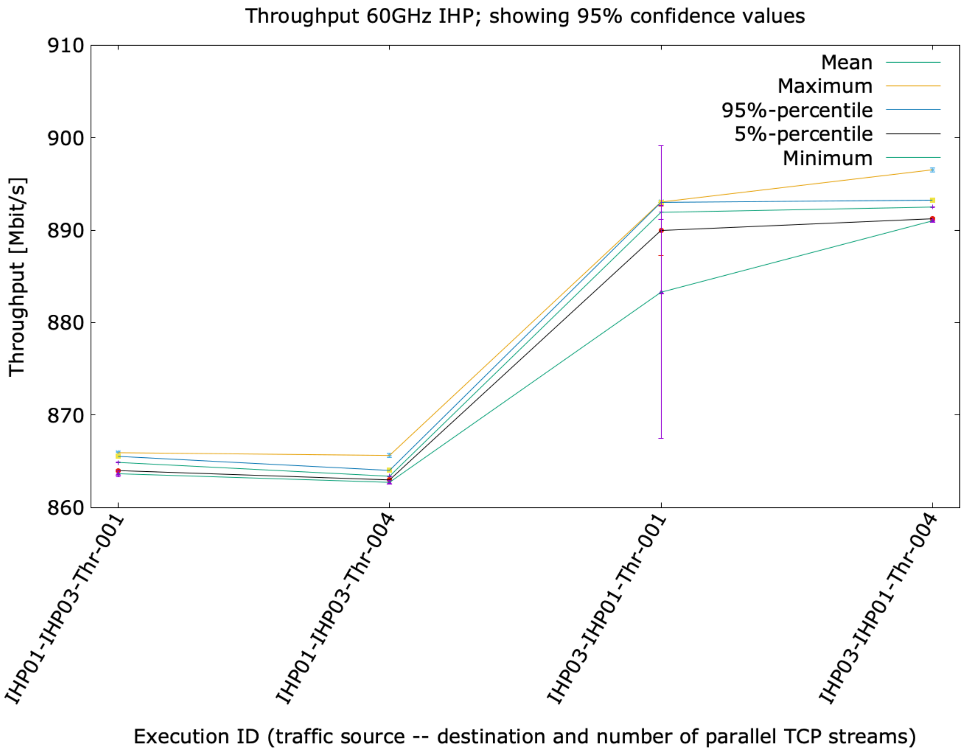


Figure 7-22 Throughput between nodes interconnected via 60 GHz IHP System (Uplink IHP01-IHP03; Downlink IHP03-iHP01)

The following table provides the detailed measurement results:Annex Table A-64.

7.2.7. Smart Grid Control Traffic

In addition to the measurements with Open5Gcore reported above, we have also evaluated its performance in the context of smart grid control traffic, examining the impact of different deployment options.

The energy sector represents undoubtedly one of the most significant challenges for 5G-enabling technologies, given its need to address a huge range of very diverse requirements to deal with across a variety of applications. There will likely be a complex web of interactions between the electricity smart grid and the communications networks including 5G. For example, in power distribution automation, switchgear interlocking between Intelligent Electronic Devices (IEDs) across bays is of paramount importance. Substation-wide switchgear interlocking must be performed in a distributed way through the exchange of messages, in the same or similar way as IEC-61850 GOOSE messages today. Hardwiring between IEDs can thus be replaced by the reliable transmission of these messages over traditional communication networks, including 5G cellular networks. In our tests, we have considered a solution, as depicted Figure 7-23, that entails tunnelling GOOSE Ethernet frames over a LTE/5G core network. We evaluated the E2E GOOSE message latency, the time that elapses between a GOOSE message being sent by an IED publisher until it is received by an IED subscriber, as our main metric. We used the 5G core latency, the latency between the eNodeB and the PGW/UPF, as a complementary metric. The latency metrics were evaluated in two deployments: a container- and a virtual machine-based deployment. According to IEC, the latency budget for GOOSE messages is 4 ms, which also served as our benchmark.

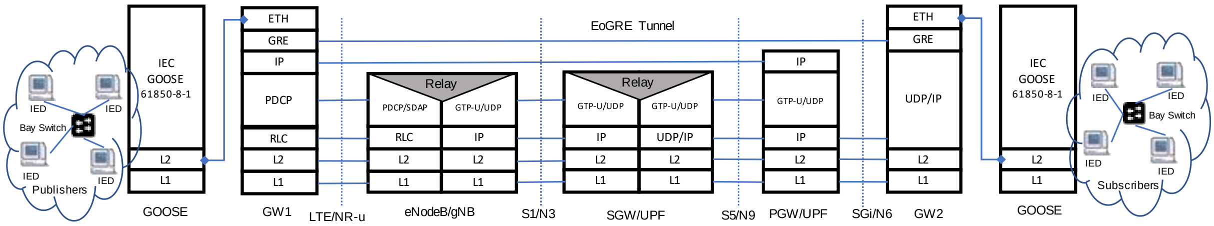


Figure 7-23 LTE/5G user plane protocol and GOOSE mapping

7.2.7.1. Test Setup

The test setup is shown in Figure 7-24. It comprised an Open5GCore server, a machine that ran the GOOSE emulation software, and two GRE-enabled GOOSE gateways. The Open5GCore server ran the Open5GCore platform (version 3), which is a standards-compliant implementation of a 5G Evolved Packet Core (EPC) system. It includes an eNodeB, a MME-SGWC-PGWC, a HSS, a SGWU-PGWU, and an INET-GW. In the VM-based deployment, these entities ran on VMs and communicated with each other using a default virtual bridge networking inside KVM. In the container-based deployment, these entities ran on Docker containers and communicated with each other using MACVLAN-based virtual networking.

The KVM and Docker machines were running Ubuntu 18.10 and were using the same specification: 8 CPUs, Intel(R) core (TM) i7-4790 CPU 3.60 GHz, and 16 GB of RAM. During the test, we also tried CPU pinning but it did not help to improve the performance. Therefore, in all tests, the CPUs were freely allocated to each VM with their corresponding number of CPUs. The GOOSE emulation machine ran the Rapid61850 software, and emulated publishing and subscribing IEDs.

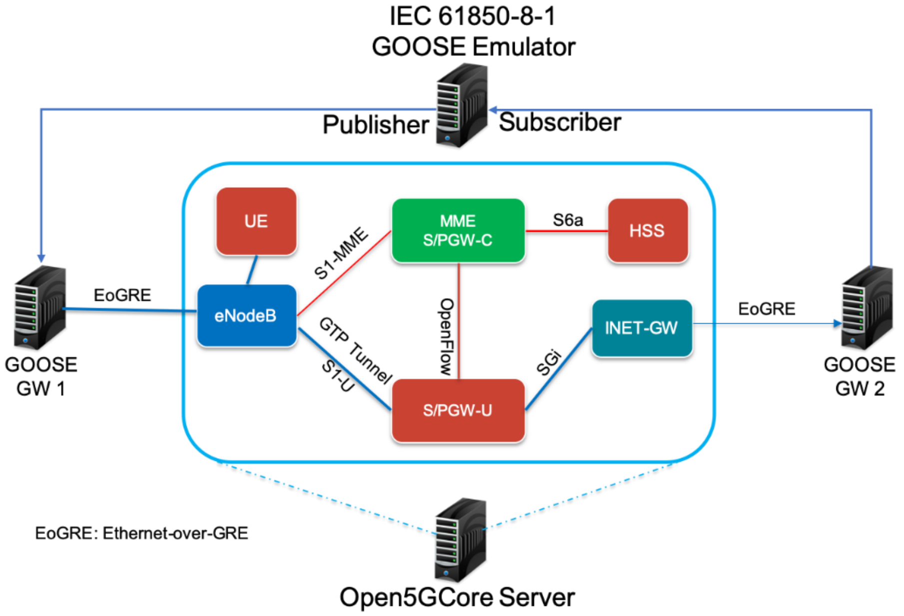


Figure 7-24 The Open5GCore test setup used in the SmartGrid tests

At the GOOSE gateways, we used Open vSwitch to setup a GRE tunneling between GOOSE GW1 and GOOSE GW2 in Figure 7-24. An Ethernet-over-GRE (EoGRE) tunneling was established between these two gateways, thus enabling the transmission of GOOSE messages encapsulated within Ethernet frames over the 5G network. The GOOSE GW1 had similar functionality as the

UE entity. It sent an attachment request to the 5G core network to be authorized, and a data-plane path was established between this gateway and the 5G Core network through the eNodeB entity, the SGWU- PGWU entity, and the INET-GW entity.

To eliminate time synchronization problems between the publisher and the subscriber, they were run on the same machine. More specifically, we used one of the interfaces of the GOOSE emulation machine for publishing GOOSE messages while listening for GOOSE messages on the other interface. These two interfaces were configured with VLAN ID 3 corresponding to pre-defined GOOSE settings. In addition, we also tuned the interrupt coalescing parameters of the network cards on each server. We set the interrupt waiting time for both the transmitting and receiving sides to be 5 µs. These are recommended for the low-latency communication.

In our tests, GOOSE frames were sent with different transmission-interval lengths, 10 ms, 20 ms, 50 ms, and with a fixed GOOSE message size of 172 bytes. Each test was repeated 30 times. The traffic was monitored through the use of Tshark. In the Open5GCore server, the GOOSE frames were captured at the ingress and egress ports (c.f. Figure 7-24). In the GOOSE emulation machine, GOOSE frames were captured at both its interfaces. The corresponding test case *TC_LAT_SmartGridControlMsgLatency_BER* is provided in the companion document "5GENESIS TEST CASES v.1.0".

7.2.7.2. Results

Figure 7-25 compiles the results from our E2E measurements. The error bars in the graphs illustrate the 95% confidence intervals.

As follows, the average E2E latency in both the KVM- and Docker-based deployments are less than 1 ms, which leaves us with a latency budget of more than 3 ms for the air interfaces between the IEDs and the GOOSE GW1 and GW2 gateways.

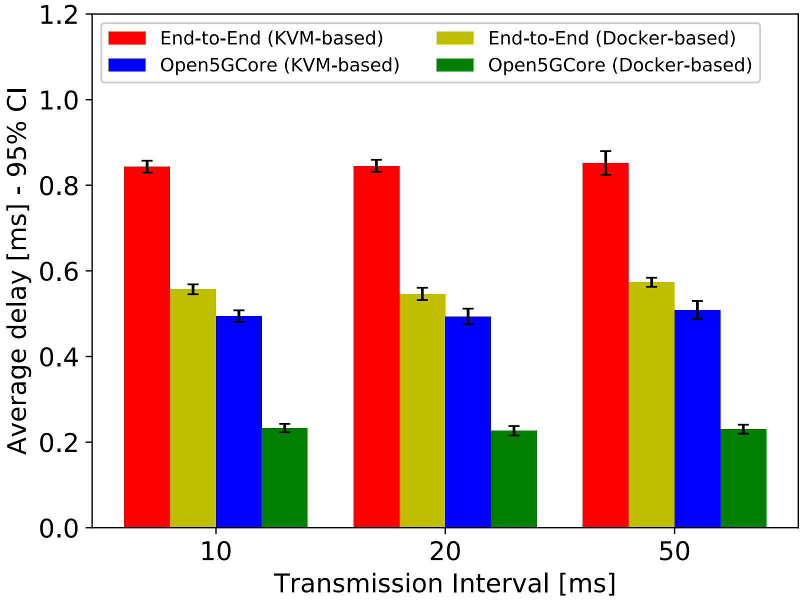


Figure 7-25 Average latency comparison between a VM- and a container-based deployment

In the Docker-based deployment, the average E2E latency at the transmission interval of 10 ms was 0.56 ms, and it was almost the same for other transmission intervals. The average 5G Core delay was about 0.23 ms. Thus, the results also show that by replacing a KVM- based

Open5GCore with a Docker-based one, the 5G Core and E2E latencies could probably be significantly reduced.

The detailed results are provided at Annex Table A-65 and Annex Table A-66.

8. SUMMARY AND CONCLUSIONS

This deliverable describes the trials and experimentation results from the second integration cycle of 5GENESIS. An upcoming version of this deliverable will describe the trials and experimentation results from third and final integration cycle (D6.3, M36).

The deliverable provides analytical results from the five 5GENESIS platforms (Málaga, Athens, Limassol, Surrey and Berlin), covering 8 main KPIs and 4 application specific validation achieved under 123 experiments performed in total. The main focus was to validate and assess the second phase integration in all platforms resulting from the deployment of 5G radio and core equipment as well as the evaluation of the Open 5GENESIS Suite [17] for automation of testing and analytics.

For all these tests, quantitative data are provided, as well as 95 % confidence intervals, based on multiple repetitions of each experiment. Specifically, for the baseline measurements the **achieved throughput of 5G ranges from 265 Mbps to 492 Mbps** subject of deployment configuration and equipment availability in each platform. Similarly, **the measured RTT ranges from 10 ms to around 40 ms for NSA deployments**.

In addition to providing measurement results, this second experimentation cycle shows **the automated experimental methodology and statistical analysis as well as the incorporation in the KPI measurement process of residual measurements** for each test case. In the next cycle, the further evolution of the platforms and the integration of additional functionalities at the coordination layer of 5GENESIS will further refine the experimentation process and the validation of KPIs of vertical cases.

9. BIBLIOGRAPHY

[1] 5GENESIS Consortium, “D2.3 Initial planning of tests and experimentation,” [Online]. Available: https://5genesis.eu/wp-content/uploads/2018/12/5GENESIS_D2.2_v1.0.pdf.

[2] 5GENESIS Consoritum, “D2.1 Requirements of the Facility,” 2018. [Online]. Available: https://5genesis.eu/wp-content/uploads/2018/11/5GENESIS_D2.1_v1.0.pdf.

[3] 5GENESIS Consortium, “D2.2 Initial overall facility design and specifications,” 2018. [Online]. Available: https://5genesis.eu/wp-content/uploads/2018/12/5GENESIS_D2.2_v1.0.pdf.

[4] 5GENESIS Consortium, “Delvierable D3.5 Monitoring and Analytics,” 2019. [Online]. Available: https://5genesis.eu/wp-content/uploads/2019/10/5GENESIS_D3.5_v1.0.pdf. [Accessed 6 2020].

[5] 5GENESIS Consortium, “The Athens Platform,” 2019. [Online]. Available: https://5genesis.eu/wp-content/uploads/2019/10/5GENESIS_D3.5_v1.0.pdf. [Accessed June 2020].

[6] 5GENESIS Consortium, “The Malaga Platform,” 2020. [Online]. Available: https://5genesis.eu/wp-content/uploads/2020/02/5GENESIS_D4.2_v1.0.pdf. [Accessed June 2020].

[7] 5GENESIS Consortium, “The Limassol Platform,” 2020. [Online]. Available: https://5genesis.eu/wp-content/uploads/2020/02/5GENESIS_D4.5_v1.0.pdf. [Accessed June 2020].

[8] 5GENESIS Consortium, “Management and Orcestration,” 2020. [Online]. Available: https://5genesis.eu/wp-content/uploads/2019/10/5GENESIS_D3.1_v1.0.pdf. [Accessed June 2020].

[9] 5GENESIS Consortium, “The Surrey Platform,” 2020. [Online]. Available: https://5genesis.eu/wp-content/uploads/2020/02/5GENESIS_D4.11_v1.0.pdf. [Accessed June 2020].

[10] 5GENESIS Consortium, “The Berlin Platform,” 2020. [Online]. Available: https://5genesis.eu/wp-content/uploads/2020/02/5GENESIS_D4.14_v1.0.pdf. [Accessed June 2020].

[11] 5GENESIS Consortium, “Deliverable D6.1 Trials and experimentation (cycle1),” 2019. [Online]. Available: http://5genesis.eu/wp-content/uploads/2019/12/5GENESIS_D6.1_v2.00.pdf. [Accessed June 2020].

[12] 3GPP, “3GPP TS 37.340 NR; Multi-connectivity; Overall description; Stage-2,” 3GPP, Dec 2018.

[13] “TRIANGLE Project, "D2.6 First Test Scenario and Test Specifications",” September 2018. [Online]. Available: https://www.triangle-project.eu/wp-content/uploads/2018/11/TRIANGLE_D2-6.pdf.

[14] 3GPP, “3GPP TS 22.179 Mission Critical Push to Talk (MCPTT); Stage 1,” 3GPP, 2019.

[15] 5GENESIS Consortium, “Deliverable D5.3 - Documentation and supporting material for 5G stakeholders,” 2020. [Online]. Available: https://5genesis.eu/wp-content/uploads/2020/07/5GENESIS-D5.3_v1.0.pdf. [Accessed June 2020].

[16] 5GENESIS Consortium, “D3.3 Slice management,” 2019. [Online]. Available: https://5genesis.eu/wp-content/uploads/2019/10/5GENESIS_D3.3_v1.0.pdf. [Accessed June 2020].

[17] 5GENESIS Consortium, “5GENESIS Github,” [Online]. Available: https://github.com/5genesis. [Accessed June 2020].

Annex A DETAILED EXPERIMENT RESULTS

A.1 Malaga Facility

A.1.1 Indoor & Outdoor e2e 5G Setup – 8.1 Full E2E 5G

A.1.1.1 Throughput

Annex Table A-1 Malaga facility Throughput KPI

Test Case ID	TC_THR_UDP			
Purpose	Measure the maximum user data rate available in different scenarios.			
Executed by	Partner:	UMA	Date:	10.04.2020
Partner(s)	UMA, NEMERGENT, ATOS, ATHONET			
Scenario	Peak reception - LOS			
Slicing configuration	-			
Components	Samsung Galaxy S10, 8.1 Deployment Setup			
Metric(s) under study	*Throughput*			
Additional tools	TAP for automated testing, VNF, iPerf, iPerf TAP plugin,			
Primary measurement results	Mean: 264.74 +/- 0.55 Mbps Standard deviation: 11.70 +/- 1.80 Mbps Median: 266.48 +/- 0.41 Mbps Min: 130.42 +/- 17.94 Mbps Max: 269.88 +/- 1.19 Mbps 25% Percentile: 265.56 +/- 0.73 Mbps 75% Percentile: 267.76 +/- 0.17 Mbps 5% Percentile: 258.78 +/- 1.46 Mbps 95% Percentile: 268.0 +/- 0.00 Mbps			
Complementary measurement results	Mean SINR (dB) 21.4; Max SINR 21.8; Min SINR 20.6 Mean RSRQ -10.8 dB Mean RSRP -58 dBm			

A.1.1.2 Round Trip Time

Annex Table A-2 Malaga facility RTT KPI

Test Case ID	TC_RTT_e2e			
General description of the test	The tests assess the average, minimum, maximum, 5% percentile and 95% percentile RTT between a UE and a VNF deployed on a single compute node in the network.			
Purpose	Measure e2e RTT			
Executed by	Partner:	UMA	Date:	11.05.2020

Partner(s)	UMA, NEMERGENT, ATOS, ATHONET
Scenario	Peak reception - LOS
Slicing configuration	
Components	Samsung Galaxy S10, 8.1 Deployment Setup
Metric(s) under study	*Round Trip Time*
Additional tools	TAP for automated testing, VNF, Ping, Ping TAP plugin,
Primary measurement results	Round Trip time [ms] Mean: 12.45 +/- 0.06 ms Standard deviation: 1.81 +/- 0.08 ms Median: 12.19 +/- 0.08 ms Min: 10.02 +/- 0.02 ms Max: 20.78 +/- 1.62 ms 25% Percentile: 11.13 +/- 0.07 ms 75% Percentile: 13.30 +/- 0.07 ms 5% Percentile: 10.237 +/- 0.04 ms 95% Percentile: 15.88 +/- 0.24 ms
Complementary measurement results	Mean SINR (dB) 21.4; Max SINR 21.8; Min SINR 20.6 Mean RSRQ -10.8 dB Mean RSRP -51 dBm

A.1.1.3 Application Specific KPI: User Experience for Content Distribution Streaming Service Use Case

Annex Table A-3 User Experience for Content Distribution Streaming Service KPI

Test Case ID	AUE/CS/001[6]			
General description of the test	Measure the user experience KPIs by the application under test while executing the feature media file playing from the Content Distribution Streaming Services use case.			
Purpose	Measure the maximum user data rate available in different scenarios.			
Executed by	Partner:	UMA	Date:	11.04.2020
Partner(s)	UMA, NEMERGENT, ATOS, ATHONET			
Scenario	Peak reception - LOS			
Slicing configuration	VNF deployed at the compute node			
Components	Samsung Galaxy S10, 8.1 Deployment Setup			
Metric(s) under study	*Time to load first media frame:* The time elapsed since the user clicks play button until the media reproduction starts			

---

[6] The test case name is taken from the list of TRIANGL Project Test Case list [13].

	Playback Cut-off: Probability that successfully started stream reproduction is ended by a cause other than the intentional termination by the user. Content Stall (s): The elapsed duration of content stalls while playing the content. Video resolution: Used video resolution.
Additional tools	TAP for automated testing, VNF, Exoplayer, Exoplayer TAP plugin,
Primary measurement results *(those included in the test case definition)*	**Time to load first media frame** 1.45 +/- 0.69 ms • **Video Resolution**ean: 3481.31 +/- 90.41 width Standard deviation: 873.69 +/- 96.84 width Median: 3635.20 +/- 292.59 width Min: 1023.84 +/- 137.87 width Max: 3840.00 +/- 0.00 width 25% Percentile: 3635.2 +/- 292.59 width 75% Percentile: 3840.0 +/- 0.00 width 5% Percentile: 1040.96 +/- 138.60 width 95% Percentile: 3840.0 +/- 0.00 width • No stalls in video reproduction • No playback cut-off
Complementary measurement results	Mean SINR 13.4 dB Mean RSRQ -2.6 dB Mean RSRP -57.5 dBm

A.1.1.4 MCPTT

Annex Table A-4 NEMERGENT MCS Access Time: 4G RAN with Athonet EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*			
General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call.			
Executed by	Partner:	UMA	Date:	05.05.2020
Partner(s)	UMA, NEM			
Scenario	Athonet Rel. 15 NSA Core with Nokia Airscale eNB with LTE band 7. The measurements are taken at the application level, in the Nemergent MCS application.			

Slicing configuration	-
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB, Athonet Rel. 15 EPC, Samsung S10 and Note10 5G UEs
Metric(s) under study	*MCPTT*
Additional tools	Logcat Android log command-line tool
Primary measurement results	**MCPTT Access time [ms]** Mean: 28.82 +/- 0.76 Standard deviation: 1.85 95% Percentile: 40.75 +/- 1.25
Complementary measurement results	n/a

Annex Table A-5 NEMERGENT MCS Access Time: 5G RAN and Athonet EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*
General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.
Purpose	Measure time from request to speak to permission granted in a MCPTT call.
Executed by	Partner: UMA Date: 23.04.2020
Partner(s)	UMA, NEM
Scenario	Athonet Rel. 15 NSA Core with Nokia Airscale eNB and gNB with LTE band 7 and 5G NR band 78. The measurements are taken at the application level, in the Nemergent MCS application.
Slicing configuration	-
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB and gNB, Athonet Rel. 15 EPC, Samsung S10 and Note10 5G UEs
Metric(s) under study	*MCPTT*
Additional tools involved	Logcat Android log command-line tool
Primary measurement results	**MCPTT Access time [ms]** Mean: 17.68 +/- 1.11 Standard deviation: 2.69 95% Percentile: 28.31 +/- 5.40

Annex Table A-6 NEMERGENT MCS Access Time: 4G RAN and Polaris NetTest EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*

General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call.			
Executed by	Partner:	UMA	Date:	04.05.2020
Partner(s)	UMA, NEM			
Scenario	Polaris NetTest Rel. 15 NSA Core with Nokia Airscale eNB with LTE band 7. The measurements are taken at the application level, in the Nemergent MCS application.			
Slicing configuration	-			
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB, Polaris NetTest Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	Logcat Android log command-line tool			
Primary measurement results	**MCPTT Access time [ms]** Mean: 27.95 +/- 0.86 Standard deviation: 2.08 95% Percentile: 41.84 +/- 1.43			

Annex Table A-7 NEMERGENT MCS Access Time: 5G RAN and Polaris NetTest EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*			
General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call.			
Executed by	Partner:	UMA	Date:	03.05.2020
Partner(s)	UMA, NEM			
Scenario	Polaris NetTest Rel. 15 NSA Core with Nokia Airscale eNB and gNB with LTE band 7 and 5G NR band 78. The measurements are taken at the application level, in the Nemergent MCS application.			
Slicing configuration	-			
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB and gNB, Polaris NetTest Rel. 15 EPC, Samsung S10 and Note10 5G UEs			

Metric(s) under study	*MCPTT*			
Additional tools	Logcat Android log command-line tool			
Primary measurement results	**MCPTT Access time [ms]** Mean: 16.72 +/- 0.78 Standard deviation: 1.89 95% Percentile: 27.61 +/- 5.05			

Annex Table A-8 NEMERGENT MCS End-to-end Access Time: 4G RAN and Athonet EPC

Test Case ID	*TC_MCPTTAccessTimeIncCallEstablishment_MAL*			
General description of the test	MCPTT end-to-end access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking, including MCPTT call establishment and possibly acknowledgment from first receiving user before voice can be transmitted.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call, including call establishment.			
Executed by	Partner:	UMA	Date:	08.05.2020
Partner(s)	UMA, NEM			
Scenario	Athonet Rel. 15 Core with Nokia Airscale eNB with LTE band 7. The measurements are taken at the application level, in the Nemergent MCS application.			
Slicing configuration	-			
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB, Athonet Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	Logcat Android log command-line tool			
Primary measurement results	**MCPTT End-to-end Access time [ms]** Mean: 137.94 +/- 3.36 Standard deviation: 8.14 95% Percentile: 190.28 +/- 4.22			

Annex Table A-9 NEMERGENT MCS End-to-end Access Time: 5G RAN with Athonet EPC

Test Case ID	*TC_MCPTTAccessTimeIncCallEstablishment_MAL*
General description of the test	MCPTT E2E access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking, including MCPTT call establishment and possibly acknowledgment from first receiving user before voice can be transmitted.

Purpose	Measure time from request to speak to permission granted in a MCPTT call, including call establishment.			
Executed by	Partner:	UMA	Date:	14.05.2020
Partner(s)	UMA, NEM			
Scenario	Athonet Rel. 15 Core with Nokia Airscale eNB and gNB with LTE band 7 and 5G NR band 78. The measurements are taken at the application level, in the Nemergent MCS application.			
Slicing configuration	-			
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB and gNB, Athonet Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	Logcat Android log command-line tool			
Primary measurement results	**MCPTT End-to-end Access time [ms]** Mean: 138.15 +/- 2.07 Standard deviation: 5.01 95% Percentile: 175.00 +/- 7.82			
Complementary measurement results	n/a			

Annex Table A-10 NEMERGENT MCS End-to-end Access Time: 4G RAN with Polaris NetTest EPC

Test Case ID	*TC_MCPTTAccessTimeIncCallEstablishment_MAL*			
General description of the test	MCPTT end-to-end access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking, including MCPTT call establishment and possibly acknowledgment from first receiving user before voice can be transmitted.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call, including call establishment.			
Executed by	Partner:	UMA	Date:	04.05.2020
Partner(s)	UMA, NEM			
Scenario	Polaris NetTest Rel. 15 Core with Nokia Airscale eNB with LTE band 7. The measurements are taken at the application level, in the Nemergent MCS application.			
Slicing configuration	-			
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB, Polaris NetTest Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	Logcat Android log command-line tool			

Primary measurement results	**MCPTT End-to-end Access time [ms]** Mean: 145.87 +/- 5.23 Standard deviation: 12.67 95% Percentile: 201.7 +/- 8.10

Annex Table A-11 NEMERGENT MCS End-to-end Access Time: 5G RAN with Polaris NetTest EPC

Test Case ID	*TC_MCPTTAccessTimeIncCallEstablishment_MAL*			
General description of the test	MCPTT E2E access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking, including MCPTT call establishment and possibly acknowledgment from first receiving user before voice can be transmitted.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call, including call establishment.			
Executed by	Partner:	UMA	Date:	20.05.2020
Partner(s)	UMA, NEM			
Scenario	Polaris NetTest Rel. 15 Core with Nokia Airscale eNB and gNB with LTE band 7 and 5G NR band 78. The measurements are taken at the application level, in the Nemergent MCS application.			
Slicing configuration	-			
Components	NEM MCS applications and MCS server VNF, Nokia Airscale eNB and gNB, Polaris NetTest Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	Logcat Android log command-line tool			
Primary measurement results	**MCPTT End-to-end Access time [ms]** Mean: 128.24 +/- 2.13 Standard deviation: 5.17 95% Percentile: 169.76 +/- 5.92			

Annex Table A-12 Airbus MCS Access Time: 4G RAN with Athonet EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*
General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.
Purpose	Measure time from request to speak to permission granted in a MCPTT call.

Executed by	Partner:	UMA	Date:	02.06.2020
Partner(s)	UMA, ADZ			
Scenario	Athonet Rel. 15 NSA Core with Nokia Airscale eNB with LTE band 7. The measurements are taken at the application level, in the Airbus Agnet MCS application.			
Slicing configuration	-			
Components	ADZ MCS applications and MCS server VNF, Nokia Airscale eNB, Athonet Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	-			
Primary measurement results	**MCPTT Access time [ms]** Mean: 40.65 +/- 1.33 Standard deviation: 3.22 95% Percentile: 55.21 +/- 3.69			

Annex Table A-13 Airbus MCS Access Time: 5G RAN with Athonet EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*			
General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call.			
Executed by	Partner:	UMA	Date:	22.05.2020
Partner(s)	UMA, ADZ			
Scenario	Athonet Rel. 15 NSA Core with Nokia Airscale eNB and gNB with LTE band 7 and 5G NR band 78. The measurements are taken at the application level, in the Airbus Agnet MCS application.			
Slicing configuration	-			
Components	ADZ MCS applications and MCS server VNF, Nokia Airscale eNB and gNB, Athonet Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	-			
Primary measurement results	**MCPTT Access time [ms]** Mean: 29.01 +/- 1.61 Standard deviation: 3.90 95% Percentile: 33.71 +/- 4.86			

Annex Table A-14 Airbus MCS Access Time: 4G RAN with Polaris NetTest EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*			
General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call.			
Executed by	Partner:	UMA	Date:	22.05.2020
Involved Partner(s)	UMA, ADZ			
Scenario	Polaris NetTest Rel. 15 NSA Core with Nokia Airscale eNB with LTE band 7. The measurements are taken at the application level, in the Airbus Agnet MCS application.			
Slicing configuration	-			
Components	ADZ MCS applications and MCS server VNF, Nokia Airscale eNB, Polaris NetTest Rel. 15 EPC, Samsung S10 and Note10 5G UEs			
Metric(s) under study	*MCPTT*			
Additional tools	-			
Primary measurement results	**MCPTT Access time [ms]** Mean: 35.96 +/- 1.09 Standard deviation: 2.64 95% Percentile: 47.81 +/- 2.09			

Annex Table A-15 Airbus MCS Access Time: 5G RAN with Polaris NetTest EPC

Test Case ID	*TC_MCPTTAccessTime_MAL*			
General description of the test	MCPTT Access time test, this test assesses the time between when an MCPTT User requests to speak and when this user gets a signal to start speaking. It does not include the MCPTT call establishment time since it measures the time previously defined when the request to speak is done during an ongoing call.			
Purpose	Measure time from request to speak to permission granted in a MCPTT call.			
Executed by	Partner:	UMA	Date:	21.05.2020
Involved Partner(s)	UMA, ADZ			
Scenario	Polaris NetTest Rel. 15 NSA Core with Nokia Airscale eNB and gNB with LTE band 7 and 5G NR band 78. The measurements are taken at the application level, in the Airbus Agnet MCS application.			
Slicing configuration	-			

Components involved	ADZ MCS applications and MCS server VNF, Nokia Airscale eNB and gNB, Polaris NetTest Rel. 15 EPC, Samsung S10 and Note10 5G UEs
Metric(s) under study	*MCPTT*
Additional tools involved	-
Primary measurement results	**MCPTT Access time [ms]** Mean: 28.10 +/- 0.63 Standard deviation: 1.52 95% Percentile: 29.29 +/- 0.55
Complementary measurement results	n/a

A.1.2 RunEl 5G RAN PHY Latency

Annex Table A-16 RunEL RAN PHY Latency

Test Case ID	TC_Lat_PHYLatency_MAL			
General description of the test	The test assesses the calculation of the average latency at PHY layer, that is, the latency of the radio interface.			
Purpose	Measure the latency for the radio interface.			
Executed by	Partner:	UMA	Date:	22.01.2020
Involved Partner(s)	UMA, REL			
Scenario	-			
Slicing configuration	-			
Components involved	RunEL setup: DRAN, RRH, UE Emulator, PC with MAC server for DRAN and UE and 2 different network interfaces (Intel X550-2).			
Metric(s) under study	*Latency*			
Additional tools involved	Iperf.			
Primary measurement results	**PHY layer latency [ms]** Mean: 1.41 +/- 0.01 Standard deviation: 0.03 95% percentile: 1.84 +/- 0.05			
Complementary measurement results	n/a			

A.2 Athens Facility

A.2.1 Amarisoft RAN – Amarisoft CN (5G.4.Option3)

A.2.1.1 Athens Facility Throughput KPI

Annex Table A-17 UDP Throughput

Test Case ID	TC_THR_Udp

General description of the test	DL Throughput of 5G NSA network			
Purpose	This test estimates the maximum user data rate in the downlink of a 5G NSA network.			
Executed by	Partner:	NCSRD	Date:	12/05/2020
Involved Partner(s)	UMA, SRL, COS			
Scenario	• One connected COTS 5G UE to the network • Radio configuration: ○ 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL ○ 4G cell: 10MHz, 2x2, Band 1, FDD			
Slicing configuration	N/A			
Components involved	• Amarisoft RAN (eNB & gNB) • Amarisoft CN Rel. 15 • Samsung A90 5G (endpoint 1) • Dell G515 (endpoint 2) • UMA iPerf & Resource Agents (Android Applications) • OpenTAP for automated testing (iPerf TAP plugin)			
Metric(s) under study	*Throughput*			
Additional tools involved	Iperf2.0.10			
Primary measurement results	Mean: 369.27 +/- 0.61 Mbps Standard deviation: 14.40 +/- 2.03 Mbps Median: 371.90 +/- 0.17 Mbps Min: 370.56 +/- 1.32 Mbps Max: 372.47 +/- 0.29 Mbps 25% Percentile: 371.40 +/- 0.63 Mbps 75% Percentile: 372.22 +/- 0.16 Mbps 5% Percentile: 368.08 +/- 2.99 Mbps 95% Percentile: 373.06 +/- 0.13 Mbps			
Complementary measurement results	Mean RSSI (dBm): -51.00 +/- 0.00 Mean RSRP (dBm): -70.48 +/- 0.20 Mean RSRQ (dB): -6.51 +/- 0.21 95th Percentile Packet Loss Rate (%): 0.95 +/- 0.78 95th Percentile Jitter (ms): 0.06 +/- 0.02			

A.2.1.2 Athens Facility Round Trip Time KPI

Annex Table A-18 Amarisoft RAN – Amarisoft EPC Round Trip Time

Test Case ID	TC_RTT_e2e			
General description of the test	Measure End-to-end RTT of 5G NSA network for packet sizes 32,64, 128, 512, 1500 bytes			
Purpose	Estimate the E2E RTT in a 5G NSA network for different packet sizes.			
Executed by	Partner:	NCSRD	Date:	14/05/2020

Involved Partner(s)	UMA, SRL, COS
Scenario	• One connected COTS 5G UE • Radio configuration: ○ 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL ○ 4G cell: 10MHz, 2x2, Band 1, FDD
Slicing configuration	N/A
Components involved	• Amarisoft RAN (eNB & gNB) • Amarisoft CN Rel. 15 • Samsung A90 5G (endpoint 1) • Dell G515 (endpoint 2) • UMA Ping & Resource Agents (Android Applications) • OpenTAP for automated testing (ping TAP plugin)
Metric(s) under study	*Round Trip Time*
Additional tools involved	ping
Primary measurement results	**E2E RTT per packet size** **32 bytes:** Mean: 31.68 +/- 0.16 ms Standard deviation: 5.97 +/- 0.10 ms Median: 31.64 +/- 0.19 ms Min: 20.97 +/- 0.68 ms Max: 44.59 +/- 1.25 ms 5% Percentile: 22.55 +/- 0.17 ms 25% Percentile: 26.64 +/- 0.18 ms 75% Percentile: 36.71 +/- 0.21 ms 95% Percentile: 40.71 +/- 0.25 ms **64 bytes:** Mean: 34.66 +/- 0.24 ms Standard deviation: 7.28 +/- 0.14 ms Median: 34.73 +/- 0.32 ms Min: 21.60 +/- 0.18 ms Max: 51.18 +/- 0.32 ms 5% Percentile: 23.06 +/- 0.19 ms 25% Percentile: 29.19 +/- 0.31 ms 75% Percentile: 39.46 +/- 0.27 ms 95% Percentile: 47.99 +/- 0.49 ms **128 bytes:** Mean: 42.04 +/- 0.18 ms Standard deviation: 6.15 +/- 0.11 ms Median: 42.14 +/- 0.23 ms

	Min: 30.75 +/- 0.73 ms Max: 57.40 +/- 1.34 ms 5% Percentile: 32.85 +/- 0.19 ms 25% Percentile: 36.91 +/- 0.22 ms 75% Percentile: 47.02 +/- 0.20 ms 95% Percentile: 50.93 +/- 0.20 ms 512 bytes: Mean: 41.75 +/- 0.15 ms Standard deviation: 5.93 +/- 0.06 ms Median: 41.81 +/- 0.18 ms Min: 31.07 +/- 0.69 ms Max: 55.60 +/- 1.00 ms 5% Percentile: 32.59 +/- 0.19 ms 25% Percentile: 36.72 +/- 0.20 ms 75% Percentile: 46.71 +/- 0.18 ms 95% Percentile: 50.72 +/- 0.17 ms 1500 bytes: Mean: 48.98 +/- 1.81 ms Standard deviation: 7.21 +/- 0.37 ms Median: 49.10 +/- 1.80 ms Min: 33.37 +/- 1.50 ms Max: 65.44 +/- 2.29 ms 5% Percentile: 36.90 +/- 1.77 ms 25% Percentile: 43.77 +/- 1.83 ms 75% Percentile: 54.55 +/- 1.94 ms 95% Percentile: 60.02 +/- 1.99 ms
Complementary measurement results	32 bytes: RSSI= -51.00+/- 0.00 dBm RSRP= -70.48 +/- 0.20dBm RSRQ=-6.51 +/- 0.21 dB Failed Ratio = 0.00 +/- 0.00 Success Ratio = 1.00 +/- 0.00 64 bytes: RSSI=-51.00 +/- 0.00 dBm RSRP= -71.80 +/- 0.22 dBm RSRQ=-6.668 +/- 0.16 dB Failed Ratio=0.00 +/- 0.00 Success Ratio=1.00 +/- 0.00 128 bytes: RSSI= -51.00 +/- 0.00 dBm

	RSRP= -69.96 +/- 0.59 dBm RSRQ= -7.36 +/- 0.39 dB Failed Ratio=0.00 +/- 0.00 Success Ratio= 1.00 +/- 0.00 **512 bytes:** RSSI= -51.00 +/- 0.00 dBm RSRP= -70.00 +/- 0.00 dBm RSRQ= -7.00 +/- 0.0 dB Failed Ratio=0.00 +/- 0.00 Success Ratio=1.00 +/- 0.00 **1500 bytes:** RSSI = -51.00 +/- 0.00 dBm RSRP=-69.00 +/- 0.00 dBm RSRQ =-6.00 +/- 0.00 dB Failed Ratio = 0.00 +/- 0.00 Success Ratio = 1.00 +/- 0.00

A.2.1.3 RTT with background traffic

Annex Table A-19 Amarisoft RAN – Amarisoft EPC Round Trip Time w bg load

Test Case ID	TC_RTT_e2eBGTraffic			
General description of the test	Measure End-to-end RTT of 5G NSA network			
Purpose	Estimate the E2E RTT in a 5G NSA network when there is concurrent traffic.			
Executed by	Partner:	NCSRD	Date:	19/05/2020
Involved Partner(s)	UMA, SRL, COS			
Scenario	• One connected COTS 5G UE • Radio configuration: ○ 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL ○ 4G cell: 10MHz, 2x2, Band 1, FDD			
Slicing configuration	N/A			
Components involved	• Amarisoft RAN (eNB & gNB) • Amarisoft Core Rel. 15 • Samsung A90 5G (endpoint 1) • Dell G515 (endpoint 2) • UMA Ping & Resource Agents (Android Applications) • OpenTAP for automated testing (ping TAP plugin)			
Metric(s) under study	*Round Trip Time*			
Additional tools involved	ping			

Primary measurement results	**End-to-end RTT [ms] - 64 bytes** Mean: 37.84 +/- 1.21 ms Standard deviation: 9.19 +/- 0.55 ms Median: 37.47 +/- 1.19 ms Min: 17.15 +/- 0.89 ms Max: 59.40 +/- 2.47 ms 5% Percentile: 24.17 +/- 0.92 ms 25% Percentile: 31.11 +/- 0.85 ms 75% Percentile: 43.65 +/- 1.33 ms 95% Percentile: 53.21 +/- 2.55 ms
Complementary measurement results	RSSI = -51.00 +/- 0.00 dBm RSRP = -71.00 +/- 0.00 dBm RSRQ - -7.00 +/- 0.00 dB Failed Ratio=0.00 +/- 0.00 Success Ratio=1.00 +/- 0.00

A.2.1.4 Athens Facility Amarisoft RAN – EPC RTT vs Radio Link Quality

Annex Table A-20 Amarisoft RAN – EPC RTT vs Radio Link Quality

Test Case ID	TC_RTT_e2eRadioLinkQuality			
General description of the test	End-to-end RTT of 5G NSA network in various cell locations with different packet sizes			
Purpose	Estimate the E2E RTT in a 5G NSA network under different RF conditions.			
Executed by	Partner:	NCSRD	Date:	01/06/2020
Involved Partner(s)	COS			
Scenario	• One connected COTS 5G UE • Radio configuration: o 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL o 4G cell: 10MHz, 2x2, Band 1, FDD • E2E RTT in Cell Edge, Mid-edge and Normal conditions			
Slicing configuration	N/A			
Components involved	• Amarisoft RAN (eNB & gNB) • Amarisoft CN Rel. 15 • Samsung A90 5G (endpoint 1) • Linux host, Ubuntu 18.04.2 LTS (endpoint 2) • MNL Metrics Tool Android Application			
Metric(s) under study	*Round Trip Time*			
Additional tools involved	ping			
Primary measurement results	**E2E RTT 64 bytes** **Cell Edge**: ~30m distance between UE and Base Station, Line of Sight			

	Mean: 32.82 +/- 0.71 ms Standard deviation: 7.22 +/- 0.25 ms Median: 32.70 +/- 0.79 ms Min: 16.67 +/- 0.23 ms Max: 47.61 +/- 1.62 ms 5% Percentile: 20.62 +/- 1.11 ms 25% Percentile: 27.94 +/- 0.90 ms 75% Percentile: 38.11 +/- 0.87 ms 95% Percentile: 44.13 +/- 0.51 ms **Mid Edge**: ~10m distance between UE and Base Station, Line of Sight Mean: 29.73 +/- 0.48 ms Standard Deviation: 8.41 +/- 1.37 ms Median: 29.48 +/- 0.44 ms Minimum: 16.89 +/- 0.21 ms Maximum: 60.50 +/- 16.41 ms 5th percentile: 18.20 +/- 0.19 ms 25th percentile: 23.67 +/- 0.50 ms 75th percentile: 34.43 +/- 0.38 ms 95th percentile: 45.11 +/- 1.08 ms **Peak conditions:** ~3m between UE and Base Station, Line of Sight Mean: 32.70 +/- 0.47 ms Standard deviation: 7.65 +/- 0.31 ms Median: 32.80 +/- 0.50 ms Min: 16.73 +/- 0.33 ms Max: 50.76 +/- 5.86 ms 5% Percentile: 19.54 +/- 0.61 ms 25% Percentile: 27.64 +/- 0.55 ms 75% Percentile: 38.28 +/- 0.68 ms 95% Percentile: 44.36 +/- 0.31 ms **E2E RTT 1500 bytes** **Cell Edge**: ~30m distance between UE and Base Station, Line of sight Mean: 45.50 +/- 0.48 ms Standard Deviation: 6.73 +/- 0.39 ms Median: 45.52 +/- 0.50 Minimum: 30.98 +/- 1.01 ms Maximum: 63.62 +/- 6.31 ms 5th percentile: 34.79 +/- 1.14 ms 25th percentile: 40.46 +/- 0.50 ms 75th percentile: 50.54 +/- 0.44 ms 95th percentile: 55.18 +/- 0.24 ms

	Mid Edge: ~10m distance between UE and Base Station, Line of sight Mean: 38.73 +/- 0.62 ms Standard Deviation: 6.73 +/- 0.23 ms Median: 38.71 +/- 0.61 ms Minimum: 26.99 +/- 0.23 ms Maximum: 58.18 +/- 3.38 ms 5th percentile: 28.53 +/- 0.46 ms 25th percentile: 33.40 +/- 0.65 ms 75th percentile: 43.69 +/- 0.60 ms 95th percentile: 49.48 +/- 1.12 ms **Peak Conditions:** ~3m distance between UE and Base Station, Line of sight Mean: 37.37 +/- 0.16 ms Standard Deviation: 6.27 +/- 0.20 ms Median: 37.43 +/- 0.23 ms Minimum: 26.66 +/- 0.23 ms Maximum: 54.51 +/- 2.69 ms 5th percentile: 27.98 +/- 0.17 ms 25th percentile: 32.07 +/- 0.16 ms 75th percentile: 42.41 +/- 0.24 ms 95th percentile: 46.24 +/- 0.21 ms
Complementary measurement results	**Cell Edge Conditions:** RSRP: -105.41 +/- 0.49 dBm RSRQ: -20.06 +/- 1.04 dB RSSI: -61.84 +/- 1.30 dBm **Mid Cell Conditions:** RSRP: -90.97 +/- 0.02 dBm RSRQ: -8.85 +/- 0.04 dB RSSI: -64.64 +/- 1.34 dBm **Peak Conditions:** RSRP= -86.00 +/- 0.00 dBm RSRQ=-7.79 +/- 0.41 dB RSSI= -61.48 +/- 1.07 dBm

A.2.1.5 Latency (one-way delay)

Annex Table A-21 One-way delay (latency)

Test Case ID	TC_Lat_e2eAppLayerLatency
General description of the test	Measure Latency (one-way delay) of 5G NSA network
Purpose	Estimate downlink/uplink latency in a 5G NSA network

Executed by	Partner:	NCSRD	Date:	27/05/2020
Involved Partner(s)	SRL, COS			
Scenario	• One connected COTS 5G UE • Radio configuration: ○ 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL ○ 4G cell: 10MHz, 2x2, Band 1, FDD			
Slicing configuration	Main Slice			
Components involved	• Amarisoft RAN (eNB & gNB) • Amarisoft Core Rel. 15 • Samsung A90 5G (endpoint 1) • Dell G515 (endpoint 2) • MNL Metrics Tool Android Application • IXIA IxChariot Traffic Generator			
Metric(s) under study	*Latency*			
Additional tools involved	MNL Metrics Tool Application			
Primary measurement results	**Downlink Latency [ms]** Mean: 21.37 +/- 0.28 ms Standard Deviation: 0.67 +/- 0.28 ms Median: 21.06 +/- 0.27 ms Minimum: 18.80 +/- 0.24 ms Maximum: 25.24 +/- 0.32 ms 5th percentile: 19.46 +/- 0.17 ms 25th percentile: 20.17 +/- 0.28 ms 75th percentile: 22.24 +/- 0.31 ms 95th percentile: 24.38 +/- 0.32 ms Estimated Error: 5.68 +/- 0.24 ms **Uplink Latency [ms]** Mean: 16.15 +/- 0.15 ms Standard Deviation: 0.37 +/- 0.15 ms Median: 16.00 +/- 0.16 ms Minimum: 11.40 +/- 0.17 ms Maximum: 20.44 +/- 0.24 ms 5th percentile: 12.23 +/- 0.20 ms 25th percentile: 14.26 +/- 0.17 ms 75th percentile: 18.28 +/- 0.22 ms 95th percentile: 20.08 +/- 0.25 ms Estimated Error: 5.78 +/- 0.18 ms			

Complementary measurement results	**Downlink Latency** Mean RSSI (dBm): -51.00 +/- 0.00 Mean RSRP (dBm): -75.00 +/- 0.00 Mean RSRQ (dB): -6.75 +/- 0.02 Jitter (ms): 0.18 +/- 0.01 Delay Factor (ms) (RTP specific): 29.14 +/- 0.17 Throughput (bps): 9986130.32 +/- 347.76 **Uplink Latency** Mean RSSI (dBm): -51.00 +/- 0.00 Mean RSRP (dBm): -75.00 +/- 0.00 Mean RSRQ (dB): -6.17 +/- 0.02 Jitter (ms): 1.01 +/- 0.01 Delay Factor (ms) (RTP specific): 38.28 +/- 0.13 Throughput (bps): 9986130.32 +/- 219.41

A.2.1.6 Service Creation Time

Annex Table A-22 Service Creation Time

Test Case ID	TC_SCT_5GConnSliceInstantiation			
General description of the test	Service Creation Time of 5G NSA end-to-end connectivity test case			
Purpose	Estimate the duration of service instantiation using the Open 5Genesis platform			
Executed by	Partner:	NCSRD	Date:	27/05/2020
Involved Partner(s)	UMA			

Scenario	• One connected COTS 5G UE
	• Radio configuration:
	  ○ 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL
	  ○ 4G cell: 10MHz, 2x2, Band 1, FDD			

Slicing configuration	Single slice connectivity service instantiation (eMBB)			
Components involved	• Amarisoft RAN (eNB & gNB) • vEPC core Rel. 14 • Samsung A90 (endpoint 1) • Dell G515 (endpoint 2)			
Metric(s) under study	*Service Creation Time*			
Additional tools involved	-			
Primary measurement results	**Service Creation Time [sec]** Mean: 63.394 sec Standard Deviation: 2.751 sec Median: 62.943 sec Minimum: 59.329 sec			

	Maximum: 72.882 sec 5th percentile: 59.498 sec 25th percentile: 61.914 sec 75th percentile: 64.270 sec 95th percentile: 71.767 sec
Complementary measurement results	**Fail Ratio** 2/25 service requests exceeded the application's timeout timers and service failed to start properly. That leads to a fail ratio of 8%

A.2.2 Amarisoft RAN & Athonet CN (5G.3.Option3)

A.2.2.1 Throughput

Annex Table A-23 Amarisoft RAN – Athonet CN Throughput

Test Case ID	TC_THR_Udp			
General description of the test	DL Throughput of 5G NSA network			
Purpose	This test estimates the maximum user data rate in the downlink of a 5G NSA network.			
Executed by	Partner:	NCSRD	Date:	28/04/2020
Partner(s)	ATH, UMA, SRL, COS			
Scenario	• One connected COTS 5G UE to the network • Radio configuration: o 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL o 4G cell: 10MHz, 2x2, Band 1, FDD			
Slicing configuration	N/A			
Components	• Amarisoft RAN (eNB & gNB) • Athonet CN Rel. 15 • Samsung A90 5G (endpoint 1) • Dell G515 (endpoint 2) • UMA iPerf & Resource Agents (Android Applications) • OpenTAP for automated testing (iPerf TAP plugin)			
Metric(s) under study	Throughput			
Additional tools	Iperf2.0.10			
Primary measurement results	Mean: 363.28 +/- 1.00 Mbps Standard deviation: 17.89 +/- 4.92 Mbps Median: 366.64 +/- 0.36 Mbps Minimum: 258.24 +/- 30.17 Mbps Maximum: 368.96 +/- 0.64 Mbps 5% Percentile: 362.75 +/- 2.41 Mbps 25% Percentile: 365.68 +/- 0.49 Mbps			

	75% Percentile: 366.88 +/- 0.18 Mbps 95% Percentile: 368.01 +/- 0.28 Mbps
Complementary measurement results	Mean RSSI (dBm): -51.00 +/- 0.00 Mean RSRP (dBm): -75.00 +/- 0.00 Mean RSRQ (dB): -7.00 +/- 0.00 95th Percentile Packet loss rate (%): 0.75 +/- 0.70 95th Percentile Jitter (ms): 0.09 +/- 0.02

A.2.2.2 Round Trip Time

Annex Table A-24 Amarisoft RAN – Athonet CN RTT

Test Case ID	TC_RTT_e2e			
General description of the test	Measure End-to-end RTT of 5G NSA network			
Purpose	Estimate the E2E RTT in a 5G NSA network when there is concurrent traffic.			
Executed by	Partner:	NCSRD	Date:	06/05/2020
Involved Partner(s)	ATH, UMA, SRL, COS			
Scenario	• One connected COTS 5G UE • Radio configuration: ○ 5G cell: 50 MHz, 2x2, Band n78, TDD, 256 QAM DL ○ 4G cell: 10MHz, 2x2, Band 1, FDD			
Slicing configuration	N/A			
Components involved	• Amarisoft RAN (eNB & gNB) • Amarisoft CN Rel. 15 • Samsung A90 5G (endpoint 1) • Dell G515 (endpoint 2) • UMA's Ping & Resource Agents (Android Applications) • OpenTAP for automated testing (ping TAP plugin)			
Metric(s) under study	Round Trip Time			
Additional tools involved	Ping			
Primary measurement results	**E2E RTT per packet size:** **32 bytes:** Mean: 32.39 +/- 0.21 ms Standard deviation: 6.05 +/- 0.15 ms Median: 32.29 +/- 0.20 ms Min: 20.95 +/- 0.65 ms Max: 46.54 +/- 1.51 ms 5% Percentile: 23.28 +/- 0.17 ms 25% Percentile: 27.08 +/- 0.22 ms 75% Percentile: 37.27 +/- 0.25 ms 95% Percentile: 41.43 +/- 0.65 ms			

	64 bytes: Mean: 32.35 +/- 0.18 ms Standard deviation: 6.44 +/- 0.16 ms Median: 32.14 +/- 0.20 ms Min: 19.88 +/- 1.10 ms Max: 48.25 +/- 1.51 ms 5% Percentile: 22.90 +/- 0.17 ms 25% Percentile: 27.10 +/- 0.25 ms 75% Percentile: 37.37 +/- 0.20 ms 95% Percentile: 41.42 +/- 0.45 ms **128 bytes:** Mean: 42.85 +/- 0.23 ms Standard deviation: 6.31 +/- 0.18 ms Median: 42.85+/- 0.28 ms Min: 27.77 +/- 1.55 ms Max: 58.60 +/- 0.89 ms 5% Percentile: 33.50 +/- 0.25 ms 25% Percentile: 37.62 +/- 0.29 ms 75% Percentile: 47.70 +/- 0.27 ms 95% Percentile: 52.14 +/- 0.54 ms **512 bytes** Mean: 42.54 +/- 0.25 ms Standard deviation: 6.29 +/- 0.34 ms Median: 42.57 +/- 0.31 ms Min: 28.67 +/- 1.45 ms Max: 55.48 +/- 1.40 ms 5% Percentile: 33.37 +/- 0.25 ms 25% Percentile: 37.57 +/- 0.27 ms 75% Percentile: 47.47 +/- 0.23 ms 95% Percentile: 51.67 +/- 0.41 ms **1500 bytes** Mean: 49.22+/- 1.86 ms Standard deviation: 7.20 +/- 0.38 ms Median: 49.22 +/- 1.83 ms Min: 32.91 +/- 1.38 ms Max: 65.83 +/- 2.45 ms 5% Percentile: 37.48 +/- 1.89 ms 25% Percentile: 43.95 +/- 1.93 ms 75% Percentile: 54.75 +/- 1.98 ms 95% Percentile: 59.99 +/- 2.06 ms
Complementary measurement results	**32 bytes** **Mean Values:** RSSI: -51.08 +/- 0.17 dBm RSRP: -74.22 +/- 0.19 dBm

	RSRQ: -6.59 +/- 0.20 dB Ping success ratio: 1.00 +/- 0.00 Ping failed ratio: 0.00 +/- 0.00 **64 bytes** **Mean Values** RSSI: -51.00 +/- 0.00 dBm RSRP: -75.00 +/- 0.03 dBm RSRQ: -7.00 +/- 0.00 dB Ping success ratio: 1.00 +/- 0.00 Ping failed ratio: 0.00 +/- 0.00 **128 bytes** **Mean Values:** RSSI: -51.00 +/- 0.00 dBm RSRP: -74.95 +/- 0.11 dBm RSRQ: -6.98 +/- 0.05 dB Ping success ratio: 1.00 +/- 0.00 Ping failed ratio: 0.00 +/- 0.00 **512 bytes** **Mean Values:** RSSI: -51.00 +/- 0.00 dBm RSRP: -75.00 +/- 0.00 dBm RSRQ: -7.00 +/- 0.00 dB Ping success ratio: 1.00 +/- 0.00 Ping failed ratio: 0.00 +/- 0.00 **1500 bytes** **Mean Values:** RSSI: -51.00 +/- 0.00 dBm RSRP: -75.00 +/- 0.00 dBm RSRQ: -7.00 +/- 0.00 dB Ping success ratio: 1.00 +/- 0.00 Ping failed ratio: 0.00 +/- 0.00

A.3 Limassol Facility

A.3.1 Edge site – Core data-center

A.3.1.1 Downlink throughput – satellite backhaul only

Annex Table A-25 Downlink throughput – satellite backhaul

Test Case ID	TC_THR_Tcp			
General description of the test	DL Throughput when terrestrial is unavailable			
Purpose	Testing the throughput of satellite backhaul			
Executed by	Partner:	SHC	Date:	14/05/2020

Involved Partner(s)	PLC, AVA, EKI, SRL, UMA
Scenario	Traffic generated from the core DC side is handled from ING multilink VM, which responds to terrestrial link outage by using satellite backhaul
Slicing configuration	N/A
Components involved	Multilink Ekinops VNF • UMA iPerf Agents • OpenTAP for automated testing (iPerf TAP plugin)
Metric(s) under study	*Throughput*
Additional tools involved	Iperf3
Test Case Statistics	Mean: 10.589 +/- 0.101 Mbps Standard deviation: 0.921 +/- 0.234 Mbps Median: 10.7 +/- 0 Mbps Min: 0.786 +/- 1.424 Mbps Max: 10.8 +/- 0.1 Mbps 25% Percentile: 10.7 +/- 0 Mbps 75% Percentile: 10.7 +/- 0 Mbps 5% Percentile: 10.7 +/- 0 Mbps 95% Percentile: 10.7 +/- 0 Mbps

A.3.1.2 Downlink throughput – terrestrial backhaul only

Annex Table A-26 Downlink throughput – terrestrial backhaul

Test Case ID	TC_THR_Tcp			
General description of the test	DL Throughput when satellite is unavailable			
Purpose	Testing the throughput of terrestrial backhaul			
Executed by	Partner:	SHC	Date:	14/05/2020
Involved Partner(s)	PLC, AVA, EKI, SRL, UMA			
Scenario	Traffic generated from the core DC side is handled from ING multilink VM, which responds to terrestrial link outage by using terrestrial backhaul			
Slicing configuration	N/A			
Components involved	• Multilink Ekinops VNF • UMA iPerf Agents • OpenTAP for automated testing (iPerf TAP plugin)			
Metric(s) under study	*Throughput*			
Additional tools involved	Iperf3			

Test Case Statistics	Mean: 17.3261 +/- 0.036 Mbps Standard deviation: 0.1062 +/- 0.0095 Mbps Median: 17.3+/- 0.0 Mbps Min: 16.0 +/- 0.1 Mbps Max: 17.5+/- 0.1 Mbps 25% Percentile: 17.3 +/- 0 Mbps 75% Percentile: 17.4 +/- 0 Mbps 5% Percentile: 17.3 +/- 0 Mbps 95% Percentile: 17.4 +/- 0 Mbps

A.3.1.3 Downlink throughput – link aggregation

Annex Table A-27 Downlink throughput – link aggregation

Test Case ID	TC_THR_Tcp			
General description of the test	DL Throughput showcasing the link aggregation mechanism on the core DC			
Purpose	Testing the total bandwidth passing through both backhaul links			
Executed by	Partner:	SHC	Date:	14/05/2020
Partner(s)	PLC, AVA, EKI, SRL, UMA			
Scenario	Traffic generated from the core DC side is handled from ING multilink VM which leverages both terrestrial and satellite links			
Components	• Multilink Ekinops VNF • UMA iPerf Agents • OpenTAP for automated testing (iPerf TAP plugin)			
Metric(s) under study	Throughput			
Additional tools	Iperf3			
Test Case Statistics	Mean: 27.7867 +/- 0.2884 Mbps Standard deviation: 0.5352 +/- 1.2022 Mbps Median: 27.804 +/- 0.496 Mbps Min: 9.69 +/- 17.51 Mbps Max: 28.9 +/- 0.4 Mbps 25% Percentile: 27.552 +/- 0.148 Mbps 75% Percentile: 28.092 +/- 0.308 Mbps 5% Percentile: 27.1688 +/- 0.3688 Mbps 95% Percentile: 28.44 +/- 0.26Mbps			

A.3.1.4 RTT – link aggregation

Annex Table A-28 RTT – link aggregation

Test Case ID	TC_RTT_e2e
General description of the test	Measure RTT over the link aggregation configuration

Purpose	Assess icmp traffic handling from multilink mechanism			
Executed by	Partner:	SHC	Date:	14/05/2020
Involved Partner(s)	PLC, AVA, EKI, SRL, UMA			
Scenario	Traffic generated from the edge cloud is handled from IUG multilink VM which splits packets on both links.			
Slicing configuration	N/A			
Components involved	• Multilink Ekinops VNF • UMA ping Agents • OpenTAP for automated testing (ping TAP plugin)			
Metric(s) under study	RTT			
Additional tools involved	ping			
Test Case Statistics	Mean: 356.2016 +/- 1.5431 ms Standard deviation: 250.1204 +/- 0.0577 ms Median: 494.06 +/- 387.06 ms Min: 102+/- 1 ms Max: 612 +/- 6 ms 25% Percentile: 105.97 +/- 0.72 ms 75% Percentile: 605.93 +/- 0.96 ms 5% Percentile: 104.438 +/- 0.562 ms 95% Percentile: 606 +/- 0 ms			

A.3.1.5 Service Creation Time – Core DC

Annex Table A-29 Service Creation Time

Test Case ID	TC_SCT_5GConnSliceInstantiation			
General description of the test	Measure slice creation time on the core DC			
Purpose	The purpose of this test is to calculate the time needed for the creation of a single-VNF slice deployed at the core data-center			
Executed by	Partner:	SHC	Date:	25/05/2020
Involved Partner(s)	NCSRD, UMA, SRL			
Scenario	The test case initiated from the Portal and ELCM, is passed to the Slice Manager which requests from OSM to create/ delete the VNF and gathers the results.			
Components involved	• Katana Slice Manager • UMA ELCM and portal • Open Source MANO			
Metric(s) under study	*Service creation time*			
Additional tools involved	Katana			
Test case statistics	Mean: 45.197664 +/- 2.064165 s			

A.3.1.6 Slice creation time - Edge node

Annex Table A-30 Slice creation time - Edge node

Test Case ID	TC_SCT_5GConnSliceInstantiation			
General description of the test	Measure slice creation time at the edge cloud			
Purpose	The purpose of this test is to calculate the time needed for the creation of a single-VNF slice deployed at the edge node			
Executed by	Partner:	SHC	Date:	25/05/2020
Involved Partner(s)	NCSRD, UMA, SRL			
Scenario	The test case initiated from the Portal and ELCM, is passed to the slice manager, which requests from OSM to create/ delete the VNF and gathers the results			
Components involved	• Katana Slice Manager • UMA ELCM and Portal • Open Source MANO			
Metric(s) under study	*Service creation time*			
Additional tools involved	Katana			
Test case statistics	Mean: 102.551440 +/- 4.285408 s			

A.3.2 5G RAN Measurements

A.3.2.1 Downlink Throughput

Annex Table A-31 Limassol 5G RAN Downlink throughput

Test Case ID	TC_THR_Tcp			
General description of the test	DL 5G Throughput			
Purpose	Testing the bandwidth over the 5G connection			
Executed by	Partner:	SHC	Date:	11/05/2020
Involved Partner(s)	SRL, UMA			
Scenario	Traffic is generated from iperf agent component over the 5G connection to the UE			
Components involved	• Galaxy A90 • UMA iPerf Agents • OpenTAP for automated testing (iPerf TAP plugin)			
Metric(s) under study	*Throughput*			
Additional tools involved	Iperf3			

Test Case Statistics	Mean: 212.72 +/- 4.00 Mbps Standard deviation: 15.44 +/- 2.04 Mbps Median: 215.75 +/- 4.10 Mbps Min: 182.14 +/- 7.57 Mbps Max: 232.57 +/- 3.47 Mbps 25% Percentile: 205.88 +/- 4.50 Mbps 75% Percentile: 222.40 +/- 3.95 Mbps 5% Percentile: 188.9 +/- 6.32 Mbps 95% Percentile: 230.20 +/- 3.57 Mbps

A.3.2.2 Uplink Throughput

Annex Table A-32 Limassol 5G RAN Uplink throughput

Test Case ID	TC_THR_Tcp			
General description of the test	UL 5G Throughput			
Purpose	Testing the bandwidth over the 5G connection			
Executed by	Partner:	SHC	Date:	11/05/2020
Involved Partner(s)	SRL, UMA			
Scenario	Traffic is generated from the UE over the 5G connection			
Slicing configuration	N/A			
Components involved	• Galaxy A90 • UMA iPerf Agents • OpenTAP for automated testing (iPerf TAP plugin)			
Metric(s) under study	*Throughput*			
Additional tools involved	Iperf3			
Test Case Statistics	Mean: 40.05 +/- 0.47 Mbps Standard deviation: 0.80 +/- 0.37 Mbps Median: 40.18 +/- 0.33 Mbps Min: 37.41 +/- 1.34 Mbps Max: 41.07 +/- 0.21 Mbps 25% Percentile: 39.68 +/- 0.73 Mbps 75% Percentile: 40.49 +/- 0.28 Mbps 5% Percentile: 39.15 +/- 1.19 Mbps 95% Percentile: 40.91 +/- 0.21 Mbps			

A.3.2.3 RTT

Annex Table A-33 Limassol E2E 5G RTT

Test Case ID	TC_RTT_e2e

General description of the test	Measure RTT			
Purpose	Measure RTT between 5G RAN and the UE			
Executed by	Partner:	SHC	Date:	11/05/2020
Involved Partner(s)	SRL, UMA			
Scenario	Ping agent component located at the edge cloud generate icmp traffic over the 5G connection to the UE			
Slicing configuration	N/A			
Components involved	• Galaxy A90 UE • UMA ping Agents • OpenTAP for automated testing (ping TAP plugin)			
Metric(s) under study	*RTT*			
Additional tools involved	ping			
Test Case Statistics	Mean: 32.26 +/- 0.65 ms Standard deviation: 7.56 +/- 0.41 ms Median: 31.96 +/- 0.82 ms Min: 17.32 +/- 0.64 ms Max: 47.42 +/- 1.26 ms 25% Percentile: 26.93 +/- 0.77 ms 75% Percentile: 37.78 +/- 1.02 ms 5% Percentile: 20.14 +/- 0.67 ms 95% Percentile: 44.34 +/- 0.92 ms			

A.4 Surrey Facility

A.4.1 5G NR (NSA) [Rel. 15]

A.4.1.1 Throughput

Annex Table A-34 Surrey 5G NR (NSA) – UDP Uplink Throughput

Test Case ID	TC-THR-Udp			
General description of the test	The test measures the average maximum user data rate available.			
Purpose	To assess the throughput in the uplink direction.			
Executed by	Partner:	UNIS	Date:	18/06/2020
Involved Partner(s)	UNIS			
Scenario	5GNR UDP			
Slicing configuration	eMBB slice			
Components involved	Rel. 15 4G Core NSA, 4G RAN (control plane), 5G RAN (user plane), 5G CPE (UE), client and server laptops			
Additional tools involved	iperf			
Primary measurement results	Mean: 119.95 +/- 0.28 Mbps Standard deviation: 2.02 +/- 0.30 Mbps Median: 120.31 +/- 0.27 Mbps Min: 111.21 +/- 3.29 Mbps Max: 125.08 +/- 0.96 Mbps 25% Percentile: 119.17 +/- 0.24 Mbps 75% Percentile: 121.15 +/- 0.37 Mbps 5% Percentile: 116.25 +/- 0.36 Mbps 95% Percentile: 122.54 +/- 0.36 Mbps 95% Percentile: 122.54230769230769 +/- 0.364535 Mbps			
Complementary measurement results	n/a			

Annex Table A-35 Surrey 5G NR (NSA) – TCP Downlink Throughput

Test Case ID	TC-THR-Tcp			
General description of the test	The test measures the average maximum user data rate available.			
Purpose	To assess the throughput in the downlink direction.			
Executed by	Partner:	UNIS	Date:	18/06/2020
Involved Partner(s)	UNIS			
Scenario	5GNR TCP			
Slicing configuration	eMBB slice			
Components involved	Rel. 15 4G Core NSA, 4G RAN (control plane), 5G RAN (user plane), 5G CPE (UE), client and server laptops			

Metric(s) under study	Throughput
Additional tools involved	iperf
Primary measurement results	Mean: 492.08 +/- 7.24 Mbps Standard deviation: 101.80 +/- 4.80 Mbps Median: 494.16 +/- 8.67 Mbps Min: 125.63 +/- 20.84 Mbps Max: 738.08 +/- 8.52 Mbps 25% Percentile: 427.87 +/- 6.40 Mbps 75% Percentile: 560.92 +/- 9.13 Mbps 5% Percentile: 332.31 +/- 13.92 Mbps 95% Percentile: 651.83 +/- 10.30 Mbps
Complementary measurement results	n/a

Annex Table A-36 Surrey 5G NR (NSA) – TCP Uplink Throughput

Test Case ID	TC-THR-Tcp			
General description of the test	The test assesses the calculation of the average maximum user data rate available.			
Purpose	To assess the throughput in the uplink direction.			
Executed by	Partner:	UNIS	Date:	16/06/2020
Involved Partner(s)	UNIS			
Scenario	5GNR iperf			
Slicing configuration	eMBB slice			
Components involved	Rel. 15 4G Core NSA, 4G RAN (control plane), 5G RAN (user plane), 5G CPE (UE), client and server laptops			
Metric(s) under study	Throughput			
Additional tools involved	iperf			
Primary measurement results	Mean: 59.34 +/- 5.29 Mbps Standard deviation: 5.95 +/- 1.16 Mbps Median: 59.72 +/- 6.07 Mbps Min: 46.38 +/- 6.00 Mbps Max: 73.40 +/- 6.30 Mbps 25% Percentile: 54.97 +/- 5.26 Mbps 75% Percentile: 63.63 +/- 5.35 Mbps 5% Percentile: 51.55 +/- 5.12 Mbps 95% Percentile: 67.52 +/- 5.04 Mbps			
Complementary measurement results	n/a			

A.4.1.2 Round Trip Time

Annex Table A-37 Surrey 5G NR (NSA) – RTT

Test Case ID	TC_RTT_e2e			
General description of the test	This test measures the mean and 95%, 5%, 25%, and 75% percentile RTT from a client (5G CPE) to a server over a 5GNR NSA Rel.15 mobile network.			
Purpose	To assess the RTT.			
Executed by	Partner:	UNIS	Date:	17/06/2020
Involved Partner(s)	UNIS			
Scenario	5GNR EMBB			
Slicing configuration	eMBB slice			
Components involved	Rel. 15 4G Core NSA, 4G RAN (control plane), 5G RAN (user plane), 5G CPE (UE), client and server laptops			
Metric(s) under study	*Round-Trip-Time*			
Additional tools involved	n/a			
Primary measurement results	Mean: 10.47 +/- 0.03 ms Standard deviation: 1.46 +/- 0.04 ms Median: 10.38 +/- 0.03 ms Min: 8.23 +/- 0.04 ms Max: 15.60 +/- 0.78 ms 25% Percentile: 9.25 +/- 0.04 ms 75% Percentile: 11.57 +/- 0.04 ms 5% Percentile: 8.49 +/- 0.01 ms 95% Percentile: 12.54 +/- 0.05 ms			
Complementary measurement results	n/a			

A.4.2 IoT use case experiments

In the tables below:

- Column 1 gives the number of packets per batch, which are tentatively sent by the board;
- Column 2 gives the number of packets per batch, which are successfully sent by the board, meaning that the sending is successfully acknowledged by the receiving party (e.g. TTN for the LoRA case);
- Column 3 gives the number of packets per batch successfully received by the server at the Surrey testbed;
- Column 4 gives the number of lost packets;
- Column 5 gives the average delay taken between the packet sending and its storage in the mySQL database, expressed in milliseconds;
- Column 6 gives the number of messages that where received with an empty payload.

A.4.2.1 Packet Delay and Packet Loss

Annex Table A-38 IoT Packet Delay HTTP POST Results

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_IoT_PacketDelayHTTP_POST_SUR</td></tr>
<tr><td>General description of the test</td><td colspan="4">The main KPI is ‘packet delays’. Under increasing load (various sampling rates), the server is able to receive, treat (decoding JSON and storing data in mySQL database) and answer all HTTP POST requests without accumulating delays. As discussed earlier we set a sampling rate of 0sec (base 1.1 sec) between each sent packet.
An *Apache* server is used at the server side and a micro-python library (urequest) is used at the board side for send HTTP POST requests.</td></tr>
<tr><td>Purpose</td><td colspan="4">Check performance (packet delay) and reliability (packet loss) of HTTP POST based transmission. Check TC_Unit_001 test case description out for more detail.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>UNIS, INF</td><td>Date:</td><td>15/06/2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">Same as above</td></tr>
<tr><td>Scenario</td><td colspan="4">See the test case description</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">n/a</td></tr>
<tr><td>Components involved</td><td colspan="4">Surrey platform: openStack, SDN switch, Apache server, mySQL + pycom/pysense board</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4">Packet delay & packet loss</td></tr>
<tr><td>Additional tools involved</td><td colspan="4">n/a</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">
<table>
<tr><th>Nbr of Packet per batch</th><th>Nbr of Sent Packets</th><th>Nbr of Rec. Packets</th><th>Lost</th><th>Average Delay</th><th>Nbr of received packets with empty payloads</th></tr>
<tr><td>1</td><td>1</td><td>1</td><td>0</td><td>121ms</td><td>0</td></tr>
<tr><td>10</td><td>10</td><td>10</td><td>0</td><td>112ms</td><td>0</td></tr>
<tr><td>100</td><td>100</td><td>100</td><td>0</td><td>123ms</td><td>0</td></tr>
<tr><td>1000</td><td>1000</td><td>1000</td><td>0</td><td>140ms</td><td>0</td></tr>
</table>
</td></tr>
<tr><td>Complementary measurement results</td><td colspan="4">n/a</td></tr>
</table>

Annex Table A-39 IoT Packet Delay MQTT results

Test Case ID	TC_IoT_PacketDelayMQTT_SUR
General description of the test	The main KPI is ‘packet delays’. Under increasing load (various sampling rates), the MQTT client at the server side is able to receive and treat (retrieving data from the queue and storing data in MySQL database) and answer all MQTT publish requests without losing data. *Mosquito* is

<table>
<tr><td></td><td colspan="4">used at the server side and a micro-python implementation of MQTT client is used at the board side (umqtt library). Sampling rate 0sec (base 1.1 sec) between each sent packet</td></tr>
<tr><td>Purpose</td><td colspan="4">Check performance (packet delay) and reliability (packet loss) of MQTT based transmission. Check TC_Unit_002 test case description out for more detail</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>UNIS, INF</td><td>Date:</td><td>15/06/2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">Same as above</td></tr>
<tr><td>Scenario</td><td colspan="4">See the test case description</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">n/a</td></tr>
<tr><td>Components involved</td><td colspan="4">Surrey platform: openStack, SDN switch, Mosquito, Apache mySQL+ Pycom/pysense board</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4">Packet delay & packet loss</td></tr>
<tr><td>Additional tools involved</td><td colspan="4">n/a</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">
<table>
<tr><td>Nbr of Packet per batch</td><td>Nbr of Sent Packets</td><td>Nbr of Rec. Packets</td><td>Nbr of Lost Packets</td><td>Average Delay</td><td>Nbr of received packets with empty payloads</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>0</td><td>92ms</td><td>0</td></tr>
<tr><td>10</td><td>10</td><td>10</td><td>0</td><td>92ms</td><td>0</td></tr>
<tr><td>100</td><td>100</td><td>100</td><td>0</td><td>94ms</td><td>0</td></tr>
<tr><td>1000</td><td>1000</td><td>1000</td><td>0</td><td>143ms</td><td>0</td></tr>
</table>
</td></tr>
<tr><td>Complementary measurement results</td><td colspan="4">n/a</td></tr>
</table>

Annex Table A-40 IoT Packet Delay CoAP Results

Test Case ID	TC_IoT_PacketDelayCoAP			
General description of the test	The main KPI is 'packet delays'. Under increasing load (various sampling rates), the CoAP server is able to receive, treat (retrieving data from the queue and storing data in mySQL database) and answer all CoAP POST requests without losing data. A custom CoAP server based on CoAPthon is used at the server side and a micro-python implementation of a CoAP client is used at the board side (microcoapy library). Sampling rate 0sec (base 1.1sec) between each sent packet.			
Purpose	Check performance (packet delay) and reliability (packet loss) of CoAP based transmission. Check TC_Unit_003 test case description out for more detail			
Executed by	Partner:	UNIS, INF	Date:	15/06/2020
Involved Partner(s)	Same as above			

<table>
<tr><td>Scenario</td><td>See the test case description</td></tr>
<tr><td>Slicing configuration</td><td>n/a</td></tr>
<tr><td>Components involved</td><td>Surrey platform: openStack, SDN switch, mySQL, Apache, in-house CoAPthon-based CoAP server implementation + Pycom/pysense board</td></tr>
<tr><td>Metric(s) under study</td><td>Packet delay & packet loss</td></tr>
<tr><td>Additional tools involved</td><td>n/a</td></tr>
<tr><td>Primary measurement results</td><td>
<table>
<tr><th>Nbr of packets per Batch</th><th>Nbr of Sent Packets[7]</th><th>Nbr of Received Packets</th><th>Nbr of Lost Packets</th><th>Average Delay</th><th>Empty payloads</th></tr>
<tr><td>1</td><td>1</td><td>1</td><td>0</td><td>107ms</td><td>0</td></tr>
<tr><td>10</td><td>10</td><td>10</td><td>0</td><td>96ms</td><td>0</td></tr>
<tr><td>100</td><td>100</td><td>90</td><td>10</td><td>95ms</td><td>0</td></tr>
<tr><td>1000</td><td>1000</td><td>539</td><td>461</td><td>103ms</td><td>0</td></tr>
</table>
</td></tr>
<tr><td>Complementary measurement results</td><td>n/a</td></tr>
</table>

Annex Table A-41 IoT Packet Delay MQTT over LoRA Results

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_IoT_PacketDelay_MQTToverLoRA_SUR</td></tr>
<tr><td>General description of the test</td><td colspan="4">The main KPI is ‘packet loss’. Under increasing dataflow (various sampling rates), the LoRA-dedicated MQTT client is able to receive, treat (retrieving data from the queue and storing data in mySQL database) and handle all IoT packets received from the MQTT Subscribe without accumulating delays. One Pycom board publishes 4 batches of IoT data (resp. 1,10,100,1000) using LoRA Send along the board-ID topic. Collecting data from the sensor and sending it as one IoT packet takes 3sec with a sampling rate equal to 0 (fastest the Pysense can do with LoRA). The average delay is calculated for each batch. We may apply additional delays if the performance are not acceptable with 0 delay (typically 2 and 5sec or more).</td></tr>
<tr><td>Purpose</td><td colspan="4">Check performance (packet delay) and reliability (packet loss) of LoRA based transmission. Check TC_Unit_004 test case description out for more detail</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>UNIS, INF</td><td>Date:</td><td>15/06/2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">Same as above</td></tr>
<tr><td>Scenario</td><td colspan="4">See the test case description</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">n/a</td></tr>
<tr><td>Components involved</td><td colspan="4">Surrey platform: openStack, SDN switch, Mosquito, Apache mySQL+ Pycom/pysense board</td></tr>
</table>

---

[7] : nbr of packets successfully sent by the Pycom board, i.e. receiving a positive acknowledge from the TTN network

<table>
<tr><td><i>(e.g. HW components, SW components)</i></td><td></td></tr>
<tr><td>Metric(s) under study
(<i>Refer to those in Section 4</i>)</td><td>Packet delay & packet loss</td></tr>
<tr><td>Additional tools involved</td><td>n/a</td></tr>
<tr><td>Primary measurement results
<i>(those included in the test case definition)</i></td><td>
Test.1
Sampling rate 2sec (base 1.1sec) between each sent packet
<table>
<tr><th>Nbr of Packet per batch</th><th>Nbr of Sent Packets</th><th>Nbr of Rec. Packets</th><th>Lost</th><th>Average Delay</th><th>Nbr of received packets with empty payloads</th></tr>
<tr><td>1</td><td>1</td><td>1</td><td>0</td><td>550ms</td><td>0</td></tr>
<tr><td>10</td><td>10</td><td>10</td><td>0</td><td>477ms</td><td>0</td></tr>
<tr><td>100</td><td>100</td><td>100</td><td>0</td><td>515ms</td><td>0</td></tr>
<tr><td>1000</td><td>1000</td><td>994</td><td>6</td><td>905ms</td><td>7</td></tr>
</table>
Test.2
Sampling rate 5sec(base 1.1sec) between each sent packet
<table>
<tr><th>Nbr of Packet per batch</th><th>Nbr of Sent Packets</th><th>Nbr of Rec. Packets</th><th>Lost</th><th>Average Delay</th><th>Nbr of received packets with empty payloads</th></tr>
<tr><td>1</td><td>1</td><td>1</td><td>0</td><td>433ms</td><td>0</td></tr>
<tr><td>10</td><td>10</td><td>9</td><td>1</td><td>448ms</td><td>0</td></tr>
<tr><td>100</td><td>100</td><td>99</td><td>1</td><td>510ms</td><td>0</td></tr>
<tr><td>1000</td><td>1000</td><td>991</td><td>9</td><td>741ms</td><td>2</td></tr>
</table>
Test.3
Sampling rate 5mn (base 1.1sec) between each sent packet
<table>
<tr><th>Nbr of Packet per batch</th><th>Nbr of Sent Packets</th><th>Nbr of Rec. Packets</th><th>Lost</th><th>Average Delay</th><th>Nbr of received packets with</th></tr>
</table>
</td></tr>
</table>

<table>
<tr>
<td></td>
<td>
<table>
<tr><td></td><td></td><td></td><td></td><td></td><td>empty payloads</td></tr>
<tr><td>300</td><td>300</td><td>297</td><td>3</td><td>534ms</td><td>2</td></tr>
</table>
<p>Considering the two first tests we can see that some packets are lost and additional few are not lost but arrive at the Surrey platform with no payload (only TTN metadata). Additionally, as mentioned earlier sending at the highest rate results in all packets being rejected by TTN. Discussing with the people at TTN it appears that there is a <i>Fair Use Policy</i> put in place by the LoRAWAN consortium, vetted by European regulation and implemented at TTN.</p>
<p>Here is a quote from a mail received from a top technical employee at TTN: <i>“Sending a message every second is not feasible with LoRAWAN unfortunately. You will be reaching duty cycle limits soon, due to European regulations, so your node will stop sending data because it is certified for the European market.”</i></p>
<p>Clearly LoRA has not been designed to handle important amount of data coming from a single board and even the size of the payload we are sending (between 70 and 75 bytes) exceeds by far the recommended size of 11 bytes. Yet what we observe as test results is not bad compared to what a strict application of the policy would have induced normally. At the end the number of lost packets remains still reasonable.</p>
<p>An additional test – a single batch of 300 packets with 5mn sampling rate- (see Test.3 above) shows much better performances in term of packet loss/empty payload even with our 70-ish bytes payload.</p>
<p>We will need to take all those conclusions into account when implementing the next test campaign involving many boards producing data in parallel. We will have in particular to implement a AttributeValueChange policy, meaning that a measurement will be sent over LoRA if and only if it exceeded the previously sent measurement by a pre-determined threshold, e.g. 10% +/- change.</p>
</td>
</tr>
<tr>
<td>Complementary measurement results</td>
<td>n/a</td>
</tr>
</table>

A.4.2.2 Narrowband IoT Energy Consumption

Annex Table A-42 NB-IoT Energy Consumption - interface synchronization delay of 0.15s

Test Case ID	TC_ENE_NBIoT_SUR

<table>
<tr><td>General description of the test</td><td colspan="4">The tests evaluate the energy consumed by an NB-IoT device during a 10-year period of operation when uplink transfers are carried out with a particular periodicity and in message bursts of particular lengths.</td></tr>
<tr><td>Purpose</td><td colspan="4">Characterize the impact of traffic pattern and network configuration on the energy consumed by an NB-IoT device.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>KAU</td><td>Date:</td><td>27.05.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">KAU, SRL</td></tr>
<tr><td>Scenario</td><td colspan="4">Air interface synchronization delay of 0.15s, C-DRX enabled.</td></tr>
<tr><td>Components</td><td colspan="4">Omnet++ simulator</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4"><i>Energy Efficiency</i>
<i>(The target KPI for Energy Efficiency is Average NB-IoT Device Energy Consumption, as defined in</i> TC_ENE_NBIoT_SUR <i>)</i></td></tr>
<tr><td>Additional tools</td><td colspan="4">N/A</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">
<table>
<tr><td colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 2 hours</td></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>8.34±0.148.34±0.14</td><td>5.01 ± 0.07</td><td>3.34 ± 0.03</td></tr>
</table>
<table>
<tr><td colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 3 hours</td></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>5.99±0.095.99±0.09</td><td>3.77±0.04</td><td>2.66±0.02</td></tr>
</table>
<table>
<tr><td colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 4 hours</td></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>4.82±0.07</td><td>3.15±0.03</td><td>2.32±0.02</td></tr>
</table>
</td></tr>
<tr><td>Complementary measurement results</td><td colspan="4">n/a</td></tr>
</table>

Annex Table A-43 NB-IoT Energy Consumption - interface synchronization delay of 0.25s

Test Case ID	TC-ENE_NBIoT_SUR

<table>
<tr><td>General description of the test</td><td colspan="4">The tests evaluate the energy consumed by an NB-IoT device during a 10-year period of operation when uplink transfers are carried out with a particular periodicity and in message bursts of particular lengths.</td></tr>
<tr><td>Purpose</td><td colspan="4">Characterize the impact of traffic pattern and network configuration on the energy consumed by an NB-IoT device.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>KAU</td><td>Date:</td><td>27.05.2020</td></tr>
<tr><td>Partner(s)</td><td colspan="4">KAU, SRL</td></tr>
<tr><td>Scenario</td><td colspan="4">Air interface synchronization delay of 0.25s, C-DRX enabled.</td></tr>
<tr><td>Components involved</td><td colspan="4">Omnet++ simulator</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4"><i>Energy Efficiency</i>
<i>(The target KPI for Energy Efficiency is Average NB-IoT Device Energy Consumption, as defined in</i> TC_ENE_NBIoT_SUR <i>)</i></td></tr>
<tr><td>Additional tools</td><td colspan="4">N/A</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 2 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>9.17±0.259.17±0.25</td><td>5.42±0.13</td><td>3.54±0.06</td></tr>
</table>
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 3 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>6.55±0.17</td><td>4.05±0.08</td><td>2.80±0.04</td></tr>
</table>
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 4 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>5.24±0.13</td><td>3.36±0.06</td><td>2.43±0.03</td></tr>
</table>
</td></tr>
<tr><td>Complementary measurement results</td><td colspan="4">n/a</td></tr>
</table>

Annex Table A-44 NB-IoT Energy Consumption - interface synchronization delay of 0.5s

Test Case ID	TC_ENE_NBIoT_SUR

<table>
<tr><td>General description of the test</td><td colspan="4">The tests evaluate the energy consumed by an NB-IoT device during a 10-year period of operation when uplink transfers are carried out with a particular periodicity and in message bursts of particular lengths.</td></tr>
<tr><td>Purpose</td><td colspan="4">Characterize the impact of traffic pattern and network configuration on the energy consumed by an NB-IoT device.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>KAU</td><td>Date:</td><td>27.05.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">KAU, SRL</td></tr>
<tr><td>Scenario</td><td colspan="4">Air interface synchronization delay of 0.5s, C-DRX enabled.</td></tr>
<tr><td>Components involved
(e.g. HW components, SW components)</td><td colspan="4">Omnet++ simulator</td></tr>
<tr><td>Metric(s) under study
(Refer to those in Section 4)</td><td colspan="4">Energy Efficiency
(The target KPI for Energy Efficiency is Average NB-IoT Device Energy Consumption, as defined in TC_ENE_NBIoT_SUR)</td></tr>
<tr><td>Additional tools involved</td><td colspan="4">N/A</td></tr>
<tr><td>Primary measurement results
(those included in the test case definition)</td><td colspan="4">
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 2 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>15.58±0.80</td><td>8.79±0.40</td><td>5.40±0.20</td></tr>
</table>
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 3 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>7.88±0.25</td><td>4.71±0.12</td><td>3.13±0.06</td></tr>
</table>
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 4 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>6.24±0.18</td><td>3.86±0.09</td><td>2.68±0.05</td></tr>
</table>
</td></tr>
<tr><td>Complementary measurement results</td><td colspan="4">n/a</td></tr>
</table>

Annex Table A-45 NB-IoT Energy Consumption - interface synchronization delay of 0.1s

Test Case ID	TC_ENE_NBIoT_SUR

<table>
<tr><td>General description of the test</td><td colspan="4">The tests evaluate the energy consumed by an NB-IoT device during a 10-year period of operation when uplink transfers are carried out with a particular periodicity and in message bursts of particular lengths.</td></tr>
<tr><td>Purpose</td><td colspan="4">Characterize the impact of traffic pattern and network configuration on the energy consumed by an NB-IoT device.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>KAU</td><td>Date:</td><td>27.05.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">KAU, SRL</td></tr>
<tr><td>Scenario</td><td colspan="4">Air interface synchronization delay of 1s, C-DRX enabled.</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">N/A</td></tr>
<tr><td>Components involved
(e.g. HW components, SW components)</td><td colspan="4">Omnet++ simulator</td></tr>
<tr><td>Metric(s) under study
(*Refer to those in Section 4*)</td><td colspan="4">*Energy Efficiency*
(The target KPI for Energy Efficiency is Average NB-IoT Device Energy Consumption, as defined in TC_ENE_NBIoT_SUR)</td></tr>
<tr><td>Additional tools involved</td><td colspan="4">N/A</td></tr>
<tr><td>Primary measurement results
(those included in the test case definition)</td><td colspan="4">
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 2 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>15.58±0.8015.58±0.80</td><td>8.79±0.40</td><td>5.40±0.20</td></tr>
</table>
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 3 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>10.82±0.53</td><td>6.29±0.26</td><td>4.03±0.13</td></tr>
</table>
<table>
<tr><th colspan="4">Average NB-IoT Device Energy Consumption [Wh], message interval of 4 hours</th></tr>
<tr><td></td><td>Burst of 1 message</td><td>Burst of 2 messages</td><td>Burst of 3 messages</td></tr>
<tr><td>Mean</td><td>8.44±0.40</td><td>5.05±0.20</td><td>3.35±0.10</td></tr>
</table>
</td></tr>
<tr><td>Complementary measurement results</td><td colspan="4">n/a</td></tr>
</table>

A.4.2.3 Narrowband IoT Coverage

Annex Table A-46 NB-IoT RAN Coverage - Deep Indoor

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_UBI_NBIoTRAN</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses NB-IoT RAN coverage and signal availability in deep indoor environments in Oslo, Norway. It comprises the results for two different NB-IoT operators and corresponding infrastructures</td></tr>
<tr><td>Purpose</td><td colspan="4">The test provides statistical indicators for the target KPI, i.e., Narrowband Reference Signal Received Power (RSRP [dBm]), and complementary metrics, i.e., Reference Signal Received Quality (RSRQ [dB]), and Signal to Interference plus Noise Ratio (SINR [dB])</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>SRL</td><td>Date:</td><td>07.2019
(over three weeks)</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">SRL, KAU</td></tr>
<tr><td>Scenario</td><td colspan="4">Deep Indoor
(14 measurement campaigns. Campaigns are defined in TC_UBI_NBIoTRAN)</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">N/A</td></tr>
<tr><td>Components involved</td><td colspan="4">Rohde&Schwarz TSMA6 spectrum scanner
Rohde&Schwarz ROMES4 software</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4">Ubiquity/Coverage
(The target KPI for Ubiquity/Coverage assessment is the NB-IoT RSRP, as defined in TC_UBI_NBIoTRAN)</td></tr>
<tr><td>Additional tools involved</td><td colspan="4">N/A</td></tr>
<tr><td>Primary measurement results
(those included in the test case definition)</td><td colspan="4"><table>
<tr><td colspan="3">NB-IoT RSRP [dBm]</td></tr>
<tr><td></td><td>Operator 1</td><td>Operator 2</td></tr>
<tr><td>Mean</td><td>-92.99</td><td>-98.64</td></tr>
<tr><td>Minimum</td><td>-108.47</td><td>-107.35</td></tr>
<tr><td>Maximum</td><td>-66.42</td><td>-85.71</td></tr>
<tr><td>Median</td><td>-92.31</td><td>-105.05</td></tr>
</table></td></tr>
</table>

5% Percentile	-105.31	-107.17
95% Percentile	-77.02	-86.23

Complementary measurement results

Complementary measurements for NB-IoT

	Operator 1		Operator 2	
	SINR [dB]	RSRQ [dB]	SINR [dB]	RSRQ [dB]
Mean	7.50	-11.05	2.21	-13.92
Min	-16.91	-23.34	-19.08	-26.02
Max	30.11	-4.48	19.34	-7.86
Median	4.53	-9.62	3.36	-10.83
5% Perc.	-13.95	-21.89	-17.01	-24.85
95% Perc.	27.55	-4.92	17.55	-7.98

Complementary measurements for LTE

	Operator 1			Operator 2		
	RSRP	SINR	RSRQ	RSRP	SINR	RSRQ
Mean	-105.99	7.78	-12.83	-108.89	1.61	-15.77
Min	-117.02	-6.34	-20.03	-140.99	-13.96	-24.34
Max	-94.16	26.23	-10.36	-92.02	18.53	-10.79
Median	-109.71	6.85	-11.23	-113.62	4.82	-14.20
5% Perc.	-113.38	-6.21	-19.75	-126.14	-10.74	-22.84

	95% Perc.	-95.53	23.08	-10.50	-92.16	14.66	-10.86

Annex Table A-47 NB-IoT RAN Coverage - Indoor

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_UBI_NBIoTRAN</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses NB-IoT RAN coverage and signal availability in indoor environments in Oslo, Norway. It comprises the results for two different NB-IoT operators and corresponding infrastructures</td></tr>
<tr><td>Purpose</td><td colspan="4">The test provides statistical indicators for the target KPI, i.e., Narrowband Reference Signal Received Power (RSRP [dBm]), and complementary metrics, i.e., Reference Signal Received Quality (RSRQ [dB]), and Signal to Interference plus Noise Ratio (SINR [dB])</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>SRL</td><td>Date:</td><td>07.2019
(over three weeks)</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">SRL, KAU</td></tr>
<tr><td>Scenario</td><td colspan="4">Indoor
(48 measurement campaigns. Campaigns are defined in TC_UBI_NBIoTRAN)</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">N/A</td></tr>
<tr><td>Components involved (e.g. HW components, SW components)</td><td colspan="4">Rohde&Schwarz TSMA6 spectrum scanner
Rohde&Schwarz ROMES4 software</td></tr>
<tr><td>Metric(s) under study (Refer to those in Section 4)</td><td colspan="4">Ubiquity/Coverage
(The target KPI for Ubiquity/Coverage assessment is the NB-IoT RSRP, as defined in TC_UBI_NBIoTRAN)</td></tr>
<tr><td>Additional tools involved</td><td colspan="4">N/A</td></tr>
</table>

Primary measurement results (those included in the test case definition)

NB-IoT RSRP [dBm]		
	Operator 1	Operator 2
Mean	-56.63	-62.94
Minimum	-72.43	-79.46
Maximum	-30.21	-29.52
Median	-55.25	-64.16
5% Percentile	-67.44	-78.61
95% Percentile	-46.58	-50.74

Complementary measurement results

Complementary measurements for NB-IoT

	Operator 1		Operator 2	
	SINR [dB]	RSRQ [dB]	SINR [dB]	RSRQ [dB]
Mean	17.51	-8.97	12.09	-10.41
Min	5.2	-12.18	0.38	-15.02
Max	26.81	-5.04	30.92	-7.68
Median	16.89	-8.84	8.97	-10.01
5% Perc.	9.32	-11.81	1.08	-14.14
95% Perc.	26.39	-6.04	28.17	-8.24

Complementary measurements for LTE

	Operator 1			Operator 2		
	RSRP	SINR	RSRQ	RSRP	SINR	RSRQ
Mean	-64.66	17.37	-10.69	-70.67	15.90	-10.40
Min	-79.77	4.59	-11.84	-86.67	4.09	-12.10
Max	-40.3	34.85	-8.1	-39.75	34.66	-7.7
Median	-64.09	17.59	-10.77	-70.23	14.69	-10.74

5% Perc.	-77.37	6.74	-11.30	-84.40	7.38	-11.31
95% Perc.	-57.05	29.73	-9.13	-58.13	25.50	-8.09

Annex Table A-48 NB-IoT RAN Coverage - Outdoor

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_UBI_NBIoTRAN</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses NB-IoT RAN coverage and signal availability in outdoor environments in Oslo, Norway. It comprises the results for two different NB-IoT operators and corresponding infrastructures</td></tr>
<tr><td>Purpose</td><td colspan="4">The test provides statistical indicators for the target KPI, i.e., Narrowband Reference Signal Received Power (RSRP [dBm]), and complementary metrics, i.e., Reference Signal Received Quality (RSRQ [dB]), and Signal to Interference plus Noise Ratio (SINR [dB])</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>SRL</td><td>Date:</td><td>07.2019
(over three weeks)</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">SRL, KAU</td></tr>
<tr><td>Scenario</td><td colspan="4">Outdoor
(22 measurement campaigns. Campaigns are defined in TC_UBI_NBIoTRAN)</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">N/A</td></tr>
<tr><td>Components involved</td><td colspan="4">Rohde&Schwarz TSMA6 spectrum scanner
Rohde&Schwarz ROMES4 software</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4">Ubiquity/Coverage
(The target KPI for Ubiquity/Coverage assessment is the NB-IoT RSRP, as defined in TC_UBI_NBIoTRAN)</td></tr>
<tr><td>Additional tools involved</td><td colspan="4">GPS antenna for location tracking</td></tr>
</table>

<table>
<tr>
<td>Primary measurement results
(those included in the test case definition)</td>
<td>
<table>
<tr><th colspan="3">NB-IoT RSRP [dBm]</th></tr>
<tr><td></td><td>Operator 1</td><td>Operator 2</td></tr>
<tr><td>Mean</td><td>-55.04</td><td>-57.85</td></tr>
<tr><td>Minimum</td><td>-68.71</td><td>-71.79</td></tr>
<tr><td>Maximum</td><td>-44.34</td><td>-51.48</td></tr>
<tr><td>Median</td><td>-54.13</td><td>-57.51</td></tr>
<tr><td>5% Percentile</td><td>-67.92</td><td>-63.09</td></tr>
<tr><td>95% Percentile</td><td>-48.67</td><td>-53.52</td></tr>
</table>
</td>
</tr>
<tr>
<td>Complementary measurement results</td>
<td>
Complementary measurements for NB-IoT
<table>
<tr><td></td><td colspan="2">Operator 1</td><td colspan="2">Operator 2</td></tr>
<tr><td></td><td>SINR [dB]</td><td>RSRQ [dB]</td><td>SINR [dB]</td><td>RSRQ [dB]</td></tr>
<tr><td>Mean</td><td>11.33</td><td>-9.59</td><td>10.29</td><td>-10.28</td></tr>
<tr><td>Min</td><td>8.16</td><td>-10.17</td><td>-8.19</td><td>-20.89</td></tr>
<tr><td>Max</td><td>17.13</td><td>-7.89</td><td>24.11</td><td>-4.51</td></tr>
<tr><td>Median</td><td>11.36</td><td>-9.73</td><td>10.07</td><td>-10.15</td></tr>
<tr><td>5% Perc.</td><td>8.89</td><td>-10.04</td><td>6.47</td><td>-10.93</td></tr>
<tr><td>95% Perc.</td><td>15.23</td><td>-7.98</td><td>23.54</td><td>-5.98</td></tr>
</table>
Complementary measurements for LTE
<table>
<tr><td></td><td colspan="3">Operator 1</td><td colspan="3">Operator 2</td></tr>
<tr><td></td><td>RSRP</td><td>SINR</td><td>RSRQ</td><td>RSRP</td><td>SINR</td><td>RSRQ</td></tr>
<tr><td>Mean</td><td>-70.67</td><td>14.48</td><td>-12.32</td><td>-71.62</td><td>13.44</td><td>-12.35</td></tr>
<tr><td>Min</td><td>-78.50</td><td>11.19</td><td>-12.99</td><td>-77.52</td><td>10.25</td><td>-13.02</td></tr>
<tr><td>Max</td><td>-55.17</td><td>29.24</td><td>-10.74</td><td>-63.83</td><td>28.95</td><td>-10.65</td></tr>
<tr><td>Median</td><td>-71.38</td><td>13.25</td><td>-12.50</td><td>-72.19</td><td>12.66</td><td>-12.49</td></tr>
</table>
</td>
</tr>
</table>

5% Perc.	-77.13	12.39	-12.91	-75.90	10.92	-12.93
95% Perc.	-58.37	24.59	-10.81	-65.58	16.52	-10.79

A.5 Berlin Facility

A.5.1 Festival of Lights 2019

A.5.1.1 Throughput

Annex Table A-49 Throughput over different network segments during the FoL-2019 field trial

Test Case ID	TC_THR_Tcp			
General description of the test	The test assesses the average Throughput between different pairs of communication end-points (VNF) deployed for the Festival of Lights field trial.			
Purpose	The test assesses the performance of the end to end connection and the of the components deployed in the field trial.			
Executed by	Partner:	FOKUS	Date:	Festival Of Lights 2019
Involved Partner(s)	FOKUS			
Scenario	An edge node with WiFi AP deployed at the HU Berlin is connected via 60GHz backhaul to a local datacenter. The local datacenter is connected to the central node in the FOKUS datacenter.			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			
Components involved	RTT Probes deployed on different points in the E2E connection, Dell R640 Server serving as HU Datacenter, Lenovo ThinkCentre serving as WiFi Client (WC), Ruckus Wireless AP as nomadic access, MetroLinQ 60GHz wireless backhaul nomadic to edge, OpenStack for virtualisation, iPerf3 for throughput traffic generation and measurement			
Metric(s) under study	*Throughput*			
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints			
Primary measurement results	**Throughput** **FOKUS Datacentre - HU Datacentre** Mean: 463,50 +/- 0,42 Mbps **Nomadic Node (Van) – HU Datacentre** Mean: 843,77 +/- 10,35 Mbps **FOKUS Datacentre – Nomadic Node (Van)** Mean: 430,26 +/- 1,01 Mbps **Nomadic Node (Van) – WiFi Client (WC)** Mean: 31,50 +/- 2,95 Mbps			
Complementary measurement results	n/a			

A.5.1.2 Round Trip Time

Annex Table A-50 RTT over different network segments during the FoL-2019 field trial

Test Case ID	TC_RTT_e2e
General description of the test	The test assesses the average, minimum, and maximum RTT between different pairs of communication endpoints (VNF) deployed for the Festival of Lights field trial.
Purpose	The test assesses the performance of the end to end connection and the of the components deployed in the field trial.
Executed by	Partner: FOKUS Date: Festival Of Lights 2019
Involved Partner(s)	FOKUS
Scenario	An edge node with WiFi AP deployed at the HU Berlin is connected via 60GHz backhaul to a local datacenter. The local datacenter is connected to the central node in the FOKUS datacenter.
Slicing configuration	An exclusive slice / software-defined network is used for the test.
Components involved	RTT Probes deployed on different points in the E2E connection, Dell R640 Server serving as HU Datacenter, Lenovo ThinkCentre serving as WiFi Client (WC), Ruckus Wireless AP as nomadic access, MetroLinQ 60GHz wireless backhaul nomadic to edge, OpenStack for virtualization, Ping for RTT measurement
Metric(s) under study	*RTT*
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints
Primary measurement results	**Round Trip time** **FOKUS – HU Datacenter 32 Byte:** Mean: 2,25 +/- 0,04 ms Min: 1,80 +/- 0,02 ms Max: 8,25 +/- 5,00 ms **FOKUS – HU Datacenter 56 Byte:** Mean: 2,61 +/- 0,04 ms Min: 1,82 +/- 0,03 ms Max: 8,62 +/- 3,63 ms **FOKUS – HU Datacenter 1400 Byte:** Mean: 2,61 +/- 0,03 ms Min: 2,22 +/- 0,03 ms Max: 6,55 +/- 1,70 ms **HU Datacenter – Nomadic Node (Van) 32 Byte:** Mean: 2,45 +/- 0,56 ms Min: 0,57 +/- 0,01 ms Max: 136,98 +/- 27,90 ms

	HU Datacenter – Nomadic Node (Van) 56 Byte: Mean: 7,70 +/- 1,03 ms Min: 0,58 +/- 0,01 ms Max: 367,22 +/- 43,51 ms **HU Datacenter – Nomadic Node (Van) 1400 Byte:** Mean: 11,29 +/- 7,07 ms Min: 0,73 +/- 0,02 ms Max: 269,16 +/- 65,74 ms **Nomadic Node (Van) – WiFi Client (WC) 32 Byte:** Mean: 8,93 +/- 1,02 ms Min: 1,74 +/- 0,11 ms Max: 120,94 +/- 31,65 ms **Nomadic Node (Van) – WiFi Client (WC) 56 Byte:** Mean: 8,11 +/- 0,36 ms Min: 1,83 +/- 0,09 ms Max: 125,02 +/- 37,78 ms **Nomadic Node (Van) – WiFi Client (WC) 1400 Byte:** Mean: 14,62 +/- 0,96 ms Min: 2,22 +/- 0,11 ms Max: 233,71 +/- 192,30 ms

A.5.1.3 360° Video Streaming

Annex Table A-51 360 Live Video Streaming Experiments

Test Case ID	TC_360LiveVideoStreamingQoE_BER			
General description of the test	The test assesses the 360° Live Video Streaming QoE KPI during the Festival of Lights 2019 in Berlin, Germany			
Purpose	The test provides initial insights on the capabilities of a 5G system towards providing 360° Live Video Streaming to massive amount of users at high levels of QoE			
Executed by	Partner:	SRL	Date:	18-20.10.2019 (two sessions: “Berlin I”, and “Berlin II”)
Involved Partner(s)	SRL, FhG, IHP			
Scenario	Outdoor nomadic network deployment			
Slicing configuration	N/A			

<table>
<tr>
<td>Components involved</td>
<td>

A custom-built JavaScript/PHP web application that captures data on web sessions, geolocation, and network connection at the client side
Bitmovin Analytics, an API-driven video analytics system that provides insight on player performance, user behavior and other aspects on the video delivery
Wireshark, which is a free and open-source packet analyzer. The application was configured to run as a background process on the web server, capturing several network and transport layer packet fields from incoming traffic

</td>
</tr>
<tr>
<td>Metric(s) under study</td>
<td>Application specific metrics
(defined in TC_360LiveVideoStreamingQoE_BER)</td>
</tr>
<tr>
<td>Additional tools involved</td>
<td>OpenTAP</td>
</tr>
<tr>
<td>Primary measurement results</td>
<td>
<p>Average time spent in different player states [sec]</p>
<table>
<tr><th></th><th>Berlin I</th><th>Berlin II</th></tr>
<tr><td>Startup</td><td>131.9</td><td>2.2</td></tr>
<tr><td>Playing</td><td>12.4</td><td>12.6</td></tr>
<tr><td>Stalling</td><td>3.8</td><td>0.7</td></tr>
</table>
<p>Total time spent on different bit rate levels while “playing” [sec]</p>
<table>
<tr><th>Bit rate [kbps]</th><th>Berlin I</th></tr>
<tr><td>1000</td><td>446.58</td></tr>
<tr><td>4000</td><td>2389.89</td></tr>
<tr><td>6000</td><td>979.86</td></tr>
<tr><td>8000</td><td>1080.18</td></tr>
</table>
<table>
<tr><th>Bit rate [kbps]</th><th>Berlin II</th></tr>
<tr><td>1100</td><td>70.89</td></tr>
<tr><td>4400</td><td>23.37</td></tr>
<tr><td>6600</td><td>765.53</td></tr>
</table>
<p>Statistics of the frequency of rebuffering events [# per minute]</p>
<table>
<tr><th></th><th>Berlin I</th><th>Berlin II</th></tr>
<tr><td>min</td><td>0.00</td><td>0.00</td></tr>
<tr><td>25% percentile</td><td>0.05</td><td>0.00</td></tr>
<tr><td>median</td><td>0.80</td><td>0.02</td></tr>
<tr><td>mean</td><td>0.81</td><td>0.11</td></tr>
<tr><td>75% percentile</td><td>1.72</td><td>0.57</td></tr>
<tr><td>max</td><td>2.24</td><td>0.70</td></tr>
</table>
<p>Statistics of Startup times per impression [sec]</p>
<table>
<tr><th></th><th>Berlin I</th><th>Berlin II</th></tr>
<tr><td>min</td><td>0.05</td><td>1.65</td></tr>
<tr><td>25% percentile</td><td>0.30</td><td>2.08</td></tr>
<tr><td>median</td><td>0.63</td><td>2.43</td></tr>
<tr><td>mean</td><td>0.58</td><td>2.47</td></tr>
<tr><td>75% percentile</td><td>1.19</td><td>2.72</td></tr>
<tr><td>max</td><td>1.65</td><td>3.53</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><td></td><td>Frequency of quality switches [# per minute]
<table>
<tr><th></th><th>Berlin I</th><th>Berlin II</th></tr>
<tr><td>min</td><td>0.00</td><td>0.00</td></tr>
<tr><td>25% percentile</td><td>0.00</td><td>0.00</td></tr>
<tr><td>median</td><td>0.00</td><td>0.07</td></tr>
<tr><td>mean</td><td>0.05</td><td>0.09</td></tr>
<tr><td>75% percentile</td><td>0.17</td><td>0.19</td></tr>
<tr><td>max</td><td>0.23</td><td>0.28</td></tr>
</table></td></tr>
<tr><td>Complementary measurement results</td><td>General statistics for “Berlin I” and “Berlin II” measurement sessions
<table>
<tr><th></th><th>Berlin I
(18/10/2019, 2 – 10 pm)</th></tr>
<tr><td>Unique impression IDs</td><td>105</td></tr>
<tr><td>Unique user IDs</td><td>28</td></tr>
<tr><td>Unique IP addresses</td><td>7</td></tr>
<tr><td>Unique user agents</td><td>11</td></tr>
<tr><td>Total play time (in seconds)</td><td>5862</td></tr>
</table>
<table>
<tr><th></th><th>Berlin II
(19-20/10/2019, 6 pm – 10 am)</th></tr>
<tr><td>Unique impression IDs</td><td>32</td></tr>
<tr><td>Unique user IDs</td><td>13</td></tr>
<tr><td>Unique IP addresses</td><td>1</td></tr>
<tr><td>Unique user agents</td><td>6</td></tr>
<tr><td>Total play time (in seconds)</td><td>3310</td></tr>
</table></td></tr>
</table>

A.5.2 OpenStack and NetApp HCI

A.5.2.1 Round Trip Time

Annex Table A-52 RTT between intra-storage-and-compute nodes within the NetApp HCI

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_RTT_e2e</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses the average, minimum, and maximum RTT between two communication endpoints (VNF) deployed on compute nodes in one data-centers.</td></tr>
<tr><td>Purpose</td><td colspan="4">The test acts as a calibration test to primarily assess the performance of the data center interconnection and the performance of the virtualization layer (SDN).</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>FOKUS</td><td>Date:</td><td>03.03.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">FOKUS</td></tr>
<tr><td>Scenario</td><td colspan="4">The communication endpoints of the service (VNFs) are within the same compute and storage unit but on different compute nodes.</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">An exclusive slice / software-defined network is used for the test.</td></tr>
<tr><td>Components involved
<i>(e.g. HW components, SW components)</i></td><td colspan="4">• VMs placed on different compute units within the same compute & storage system
• Compute and storage system (NetApp HCI)</td></tr>
</table>

Metric(s) under study (*Refer to those in Section 4*)	*RTT*
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints
Primary measurement results *(those included in the test case definition)*	**Round Trip time** **32 Byte:** Mean: 0.20 +/- 0.00 ms Standard deviation: 0.04 +/- 0.01 ms Median: 0.20 +/- 0.00 ms Min: 0.13 +/- 0.01 ms Max: 0.39 +/- 0.09 ms 25% Percentile: 0.18 +/- 0.00 ms 75% Percentile: 0.22 +/- 0.00 ms 5% Percentile: 0.16 +/- 0.00 ms 95% Percentile: 0.26 +/- 0.00 ms **56 Byte:** Mean: 0.20 +/- 0.00 ms Standard deviation: 0.04 +/- 0.01 ms Median: 0.20 +/- 0.00 ms Min: 0.13 +/- 0.01 ms Max: 0.42 +/- 0.08 ms 25% Percentile: 0.18 +/- 0.00 ms 75% Percentile: 0.22 +/- 0.00 ms 5% Percentile: 0.16 +/- 0.00 ms 95% Percentile: 0.25 +/- 0.00 ms **1400 Byte:** Mean: 0.20 +/- 0.00 ms Standard deviation: 0.03 +/- 0.00 ms Median: 0.20 +/- 0.00 ms Min: 0.13 +/- 0.01 ms Max: 0.32 +/- 0.03 ms 25% Percentile: 0.18 +/- 0.00 ms 75% Percentile: 0.21 +/- 0.00 ms 5% Percentile: 0.16 +/- 0.00 ms 95% Percentile: 0.25 +/- 0.00 ms

Annex Table A-53 RTT across the central DellSwitch (between NetApp HCI and ThinkCenter)

Test Case ID	TC_RTT_e2e
General description of the test	The test assesses the average, minimum, and maximum RTT between two communication end-points (VNF) deployed different compute nodes connected via a local switch.
Purpose	The test acts as a calibration test to primarily assess the performance of the data center interconnection and the performance of the virtualization layer (SDN).

Executed by	Partner:	FOKUS	Date:	03.03.2020
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints of the service (VNFs) are within the same compute and storage unit but on different compute nodes.			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			
Components involved *(e.g. HW components, SW components)*	• VMs placed on different compute units • Compute and storage system (NetApp HCI) • ThinkCenter • Dell Switch			
Metric(s) under study (*Refer to those in Section 4*)	*RTT*			
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication end-points			
Primary measurement results *(those included in the test case definition)*	**Round Trip time** **32 Byte:** Mean: 1.13 +/- 0.17 ms Standard deviation: 2.41 +/- 1.81 ms Median: 0.87 +/- 0.00 ms Min: 0.78 +/- 0.01 ms Max: 26.70 +/- 19.83 ms 25% Percentile: 0.84 +/- 0.00 ms 75% Percentile: 0.92 +/- 0.01 ms 5% Percentile: 0.81 +/- 0.00 ms 95% Percentile: 1.18 +/- 0.06 ms **56 Byte:** Mean: 1.08 +/- 0.14 ms Standard deviation: 1.73 +/- 1.25 ms Median: 0.87 +/- 0.00 ms Min: 0.77 +/- 0.01 ms Max: 18.86 +/- 13.17 ms 25% Percentile: 0.85 +/- 0.00 ms 75% Percentile: 0.91 +/- 0.01 ms 5% Percentile: 0.81 +/- 0.00 ms 95% Percentile: 1.17 +/- 0.05 ms **1400 Byte:** Mean: 1.50 +/- 0.37 ms Standard deviation: 4.72 +/- 3.63 ms Median: 0.98 +/- 0.01 ms Min: 0.87 +/- 0.01 ms Max: 50.78 +/- 39.48 ms 25% Percentile: 0.96 +/- 0.00 ms 75% Percentile: 1.02 +/- 0.01 ms 5% Percentile: 0.91 +/- 0.01 ms			

	95% Percentile: 1.39 +/- 0.08 ms

A.5.2.2 Throughput

Annex Table A-54 Throughput between VMs deployed within the NetApp HCI

Test Case ID	TC_THR_Udp			
General description of the test	The test assesses the average Throughput between two communication end-points (VNF) deployed on one compute and storage system featuring several compute nodes.			
Purpose	The test acts as a calibration test to primarily assess the performance of the compute node and the performance of the virtualization layer (SDN) as the entire communication resides within a single compute node, i.e. never leaves the node and hence never transverses a physical network interface, nor intermediate physical switches.			
Executed by	Partner:	FOKUS	Date:	04.03.2020
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints of the service (VNFs) are both placed within one compute and storage node (NetApp HCI). A single data stream is established between the two units.			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			
Components involved *(e.g. HW components, SW components)*	Two VMs acting as measurement endpoints, both placed on the same compute unit within the compute & storage system (NetApp HCI)			
Metric(s) under study	*Throughput*			
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication end-points			
Primary measurement results	C0U0-C0U1-TC-Thr-001 Mean: 19054.80 +/- 331.32 Mbps Standard deviation: 1049.29 +/- 229.33 Mbps Median: 19162.60 +/- 305.08 Mbps Min: 15619.16 +/- 1681.12 Mbps Max: 25539.84 +/- 1505.72 Mbps 25% Percentile: 18624.11 +/- 310.42 Mbps 75% Percentile: 19622.01 +/- 304.26 Mbps 5% Percentile: 17482.09 +/- 620.37 Mbps 95% Percentile: 20181.39 +/- 309.50 Mbps			
Complementary measurement results	n/a			

Annex Table A-55 Throughput between two VMs within the NetApp HCI

Test Case ID	TC_THR_Tcp

General description of the test	The test assesses the average Throughput between two communication end-points (VNF) deployed on one compute and storage system featuring several compute nodes.			
Purpose	The test acts as a calibration test to primarily assess the performance of the compute node and the performance of the virtualization layer (SDN) as the entire communication resides within a single compute node, i.e. never leaves the node and hence never transverses a physical network interface, nor intermediate physical switches.			
Executed by	Partner:	FOKUS	Date:	04.03.2020
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints of the service (VNFs) are both placed within one compute and storage node (NetApp HCI). Four parallel data streams are established between the two units.			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			
Components involved	Two VMs acting as measurement endpoints, both placed on the same compute unit within the compute & storage system (NetApp HCI).			
Metric(s) under study	*Throughput*			
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints			
Primary measurement results	C0U0-C0U1-TC_THR_Udp Mean: 28658.97 +/- 184.62 Mbps Standard deviation: 861.29 +/- 61.61 Mbps Median: 28653.74 +/- 177.33 Mbps Min: 26340.68 +/- 206.36 Mbps Max: 30898.28 +/- 330.95 Mbps 25% Percentile: 28067.19 +/- 166.86 Mbps 75% Percentile: 29261.59 +/- 213.67 Mbps 5% Percentile: 27229.92 +/- 138.48 Mbps 95% Percentile: 30026.44 +/- 269.38 Mbps			
General description of the test	The test assesses the average Throughput between two communication end-points (VNF) deployed on one compute and storage system featuring several compute nodes.			
Purpose	The test acts as a calibration test to primarily assess the performance of the compute node and the performance of the virtualization layer (SDN) as the entire communication resides within a single compute node, i.e. never leaves the node and hence never transverses a physical network interface, nor intermediate physical switches.			
Executed by	Partner:	FOKUS	Date:	04.03.2020
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints of the service (VNFs) are both placed on different compute nodes with the compute and storage node (NetApp HCI). A single data stream is established between the two units.			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			

<table>
<tr><td>Components involved</td><td colspan="4">Two VMs acting as measurement endpoints, placed on different compute nodes within the compute & storage system (NetApp HCI), i.e. C0U0-C1U0-THR1.</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4"><i>Throughput</i></td></tr>
<tr><td>Additional tools involved</td><td colspan="4">TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">C0U0-C1U0-TC_THR_Tcp
Mean: 13784.60 +/- 68.83 Mbps
Standard deviation: 521.03 +/- 20.71 Mbps
Median: 13822.68 +/- 70.57 Mbps
Min: 12001.60 +/- 292.70 Mbps
Max: 16290.28 +/- 404.08 Mbps
25% Percentile: 13487.58 +/- 67.17 Mbps
75% Percentile: 14130.90 +/- 72.49 Mbps
5% Percentile: 12925.02 +/- 73.80 Mbps
95% Percentile: 14463.99 +/- 72.36 Mbps</td></tr>
<tr><td>Test Case ID</td><td colspan="4">TC_THR_Udp</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses the average Throughput between two communication end-points (VNF) deployed on one compute and storage system featuring several compute nodes.</td></tr>
<tr><td>Purpose</td><td colspan="4">The test acts as a calibration test to primarily assess the performance of the compute node and the performance of the virtualization layer (SDN) as the entire communication resides within a single compute node, i.e. never leaves the node and hence never transverses a physical network interface, nor intermediate physical switches.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>FOKUS</td><td>Date:</td><td>04.03.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">FOKUS</td></tr>
<tr><td>Scenario</td><td colspan="4">The communication endpoints of the service (VNFs) are both placed within one compute and storage node (NetApp HCI). Four parallel data stream is established between the two units.</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">An exclusive slice / software-defined network is used for the test.</td></tr>
<tr><td>Components involved</td><td colspan="4">Two VMs acting as measurement endpoints, both placed different compute nodes within the compute & storage system (NetApp HCI), i.e. C0U0-C1U0-THR4.</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4"><i>Throughput</i></td></tr>
<tr><td>Additional tools involved</td><td colspan="4">TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication end-points</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">C0U0-C1U0-TC_THR_Udp
Mean: 16601.00 +/- 47.90 Mbps
Standard deviation: 298.96 +/- 14.50 Mbps
Median: 16645.50 +/- 51.31 Mbps
Min: 15157.56 +/- 263.81 Mbps
Max: 17134.56 +/- 46.62 Mbps
25% Percentile: 16455.19 +/- 50.36 Mbps
75% Percentile: 16804.73 +/- 47.95 Mbps</td></tr>
</table>

	5% Percentile: 16089.021999999999 +/- 62.699108 Mbps 95% Percentile: 16984.392 +/- 47.807415 Mbps

Annex Table A-56 Throughput between NetApp HCI and ThinkCenter over DellSwitch

Test Case ID	TC_THR_Tcp			
General description of the test	The test assesses the average Throughput between two communication end-points (VNF) interconnected over the Dell switch.			
Purpose	The test acts as a calibration test to primarily assess the performance of the involved system components.			
Executed by	Partner:	FOKUS	Date:	04.03.2020
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints of the service (VNFs) are placed at the NetApp HCI compute and storage unit and on a ThinkCenter, both interconnected via a Dell Switch. The test employs a single data stream between the components. Results are evaluated for the up- and down-link direction, i.e. Switch-C0U0-Thr-01 and C0U0-Switch-Thr-01.			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			
Components involved	• Two VMs acting as measurement endpoints, • ThinkCenter • NetApp HCI • Dell Switch			
Metric(s) under study	*Throughput*			
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication end-points			
Primary measurement results	**Switch-C0U0-THR_Tcp** Mean: 941.96 +/- 0.01 Mbps Standard deviation: 3.92 +/- 0.13 Mbps Median: 941.12 +/- 0.14 Mbps Min: 930.64 +/- 0.59 Mbps Max: 970.32 +/- 1.09 Mbps 25% Percentile: 941.00 +/- 0.00 Mbps 75% Percentile: 942.15 +/- 0.17 Mbps 5% Percentile: 937.71 +/- 0.79 Mbps 95% Percentile: 949.94 +/- 0.70 Mbps **C0U0-Switch-THR-Tcp** Mean: 934.52 +/- 0.03 Mbps Standard deviation: 5.45 +/- 0.09 Mbps Median: 931.90 +/- 0.50 Mbps Min: 926.28 +/- 1.13 Mbps Max: 949.16 +/- 0.80 Mbps 25% Percentile: 929.96 +/- 0.08 Mbps 75% Percentile: 940.81 +/- 0.27 Mbps 5% Percentile: 928.88 +/- 0.18 Mbps			

	95% Percentile: 942.804 +/- 0.203700 Mbps

Annex Table A-57 Throughput between NetApp HCI and ThinkCenter over DellSwitch

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_THR_Udp</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses the average Throughput between two communication end-points (VNF) interconnected over the Dell switch.</td></tr>
<tr><td>Purpose</td><td colspan="4">The test acts as a calibration test to primarily assess the performance of the involved system components.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>FOKUS</td><td>Date:</td><td>04.03.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">FOKUS</td></tr>
<tr><td>Scenario</td><td colspan="4">The communication endpoints of the service (VNFs) are placed at the NetApp HCI compute and storage unit and on a ThinkCenter, both interconnected via a Dell Switch. The test employs a four parallel data stream between the components. Results are evaluated for the up- and down-link direction, i.e. Switch-C0U0-Thr-04 and C0U0-Switch-Thr-04.</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">An exclusive slice / software-defined network is used for the test.</td></tr>
<tr><td>Components involved</td><td colspan="4">• Two VMs acting as measurement endpoints,
• ThinkCenter
• NetApp HCI
• Dell Switch</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4"><i>Throughput</i></td></tr>
<tr><td>Additional tools involved</td><td colspan="4">TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication end-points</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">C0U0-SWITCH-TC-Thr-Udp
Mean: 934.72 +/- 0.03 Mbps
Standard deviation: 6.61 +/- 0.34 Mbps
Median: 934.22 +/- 0.25 Mbps
Min: 919.56 +/- 1.53 Mbps
Max: 975.48 +/- 0.97 Mbps
25% Percentile: 930.42 +/- 0.29 Mbps
75% Percentile: 938.46 +/- 0.53 Mbps
5% Percentile: 925.43 +/- 0.67 Mbps
95% Percentile: 944.65 +/- 0.774671 Mbps
SWITCH-C0U0-TC-Thr-Udp
Mean: 941.79 +/- 0.45 Mbps
Standard deviation: 14.40 +/- 5.10 Mbps
Median: 942.08 +/- 0.25 Mbps
Min: 860.88 +/- 80.70 Mbps
Max: 1022.64 +/- 2.67 Mbps
25% Percentile: 935.49 +/- 0.44 Mbps
75% Percentile: 947.87 +/- 0.35 Mbps
5% Percentile: 927.09 +/- 0.68 Mbps
95% Percentile: 955.86 +/- 0.62 Mbps</td></tr>
</table>

A.5.3 Packet Core

A.5.3.1 Round Trip Time

Annex Table A-58 E2E 5G-SA end-to-end RTT (using the Open5GCore UE/NB Emulator)

Test Case ID	TC_RTT_e2e			
General description of the test	The test assesses the average, minimum, and maximum RTT between two communication end-points, namely the UE and an additional host connected to the core network. The UE and gNB are emulated via the Open5GCore UE/gNB Emulator.			
Purpose	The test acts as a calibration test to primarily assess the performance of the LTE link.			
Executed by	Partner:	FOKUS	Date:	10.07.2019
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints are the UE and the additional host.			
Slicing configuration	An exclusive slice is used for the test.			
Components involved	• UE/gNB Emulator • Open5GCore • VM as communicatioin endpoint at the network side			
Metric(s) under study	*RTT*			
Additional tools involved	TAP for automated testing			
Primary measurement results	**Round Trip time** **R640_LXC_vUE_v5GCore-32** Mean: 0.59 +/- 0.01 ms Standard deviation: 0.04 +/- 0.00 ms Median: 0.59 +/- 0.01 ms Min: 0.46 +/- 0.01 ms Max: 0.68 +/- 0.03 ms 25% Percentile: 0.57 +/- 0.01 ms 75% Percentile: 0.61 +/- 0.00 ms 5% Percentile: 0.52 +/- 0.01 ms 95% Percentile: 0.64 +/- 0.00 ms **R640_LXC_vUE_v5GCore-56** Mean: 0.58 +/- 0.00 ms Standard deviation: 0.04 +/- 0.00 ms Median: 0.59 +/- 0.01 ms Min: 0.45 +/- 0.01 ms Max: 0.71 +/- 0.08 ms 25% Percentile: 0.56 +/- 0.01 ms 75% Percentile: 0.61 +/- 0.01 ms 5% Percentile: 0.52 +/- 0.01 ms 95% Percentile: 0.646625 +/- 0.00 ms **R640_LXC_vUE_v5GCore-1300** Mean: 0.59 +/- 0.00 ms			

	Standard deviation: 0.04 +/- 0.01 ms Median: 0.60 +/- 0.01 ms Min: 0.46 +/- 0.02 ms Max: 0.74 +/- 0.07 ms 25% Percentile: 0.57 +/- 0.01 ms 75% Percentile: 0.61 +/- 0.01 ms 5% Percentile: 0.53 +/- 0.01 ms 95% Percentile: 0.65 +/- 0.01 ms **R640_LXC_vUE_v5GCore-1400** Mean: 0.60 +/- 0.07 ms Standard deviation: 0.04 +/- 0.00 ms Median: 0.61 +/- 0.01 ms Min: 0.48 +/- 0.01 ms Max: 0.73 +/- 0.04 ms 25% Percentile: 0.58 +/- 0.01 ms 75% Percentile: 0.62 +/- 0.01 ms 5% Percentile: 0.53 +/- 0.01 ms 95% Percentile: 0.66 +/- 0.01 ms
Complementary measurement results	n/a

A.5.3.2 Throughput

Annex Table A-59 5G-SA end-to-end throughput (using the Open5GCore UE/NB Emulator)

<table>
<tr><td>Test Case ID</td><td colspan="4">TC_THR_Tcp / TC_THR_Udp</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses the average throughput between two communication end-points, namely the UE and an additional host connected to the core network</td></tr>
<tr><td>Purpose</td><td colspan="4">The test acts as a calibration test to primarily assess the downlink performance of the 5G Core Network link.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>FOKUS</td><td>Date:</td><td>18.03.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">FOKUS</td></tr>
<tr><td>Scenario</td><td colspan="4">The communication endpoints are placed within the UE and the additional host.</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">An exclusive slice is used for the test.</td></tr>
<tr><td>Components involved</td><td colspan="4">• Measurement probe placed at the UE emulator
• UE/gNB Emulator
• Open5GCore
• NetApp HCI</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4">*Throughput*</td></tr>
<tr><td>Additional tools involved</td><td colspan="4"></td></tr>
<tr><td>Primary measurement results</td><td colspan="4">**Throughput V5G-C0U0-5GCore-vUE-TC-Thr-Tcp**
Mean: 753.69 +/- 4.86 Mbps
Standard deviation: 67.21 +/- 2.32 Mbps
Median: 753.70 +/- 9.92 Mbps</td></tr>
</table>

	Min: 503.72 +/- 26.00 Mbps Max: 873.40 +/- 5.78 Mbps 25% Percentile: 702.06 +/- 7.54 Mbps 75% Percentile: 809.98 +/- 5.62 Mbps 5% Percentile: 654.56 +/- 7.31 Mbps 95% Percentile: 848.76 +/- 4.57 Mbps **Throughput V5G-C0U0-5GCore-vUE-TC-Thr-Udp** Mean: 711.08 +/- 4.82 Mbps Standard deviation: 55.09 +/- 1.89 Mbps Median: 716.04 +/- 5.46 Mbps Min: 470.68 +/- 22.90 Mbps Max: 815.84 +/- 5.39 Mbps 25% Percentile: 671.23 +/- 6.21 Mbps 75% Percentile: 752.61 +/- 5.77 Mbps 5% Percentile: 625.41 +/- 6.69 Mbps 95% Percentile: 789.93 +/- 4.85 Mbps

A.5.3.3 User Density

Annex Table A-60 User Density: Number of consecutively supported user (Open5GCore)

Test Case ID	TC_DEN_MaxRegisteredUE_BER_001, TC_DEN_MaxRegisteredUE_BER_002				
General description of the test	The test assesses the maximum number of simulated users that can be served by a 5G Core Network instance simultaneously.				
Purpose	Tests serves to evaluate the 5G Core Network performance with regards to user density.				
Executed by	Partner:	FOKUS	Date:	Open5GCore IxLoad	09-15.04.2020 27-30.04.2020
Involved Partner(s)	FOKUS				
Scenario	A 5G UE and gNB emulation tool is connected to a5G Core Network, to simulate a high number of simultaneous user connections.				
Slicing configuration	An exclusive slice is used for the test.				
Components involved	• UE/gNB Emulator (Open5GCore-BT, IxLoad) • Open5GCore • OpenStack				
Metric(s) under study	*User Density* *Maximum Attached users*				
Additional tools involved					
Primary measurement results	**User Density Maximum Simultaneously Served Users** **Open5GCore BT TC_DEN_MaxRegisteredUE_BER_001** Mean: 2777,78 +/- 390,49 StdDev: 3008,24 **IxLoadCore TC_DEN_MaxRegisteredUE_BER_002** Mean: 3329,71 +/- 852,86				

Complementary measurement results	n/a

A.5.4 60GHz MetroLinq

A.5.4.1 Round Trip Time

Annex Table A-61 RTT between nodes interconnected via 60 GHz Metrolinq System

Test Case ID	TC_RTT_e2e			
General description of the test	The test assesses the average, minimum, and maximum RTT between two communication endpoints (VNF) deployed on both sides of a 60 GHz backhaul.			
Purpose	The test acts as a calibration test to primarily assess the performance of the intermediate 60GHz backhaul link.			
Executed by	Partner:	FOKUS	Date:	06.03.2020
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints of the service (VNFs), i.e. the virtual instruments, are placed within the FOKUS03 and the NetApp HCI, both interconnected via a DellSwitch			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			
Components involved	• ThinkCenter • DellSwitch • NetApp HCI • The connecting 60 GHz backhaul is realized by the MetroLinq nodes.			
Metric(s) under study	*RTT*			
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints			
Primary measurement results	Round Trip Time **60GHz-C0U0-32** Mean: 1.35 +/- 0.25 ms Standard deviation: 3.46 +/- 2.62 ms Median: 0.98 +/- 0.00 ms Min: 0.82 +/- 0.01 ms Max: 38.50 +/- 28.89 ms 25% Percentile: 0.93 +/- 0.01 ms 75% Percentile: 1.03 +/- 0.00 ms 5% Percentile: 0.88 +/- 0.01 ms 95% Percentile: 1.39 +/- 0.08 ms **60GHz-C0U0-56** Mean: 1.61 +/- 0.40 ms Standard deviation: 5.97 +/- 4.23 ms Median: 0.98 +/- 0.01 ms Min: 0.82 +/- 0.01 ms Max: 64.86 +/- 46.40 ms 25% Percentile: 0.93 +/- 0.01 ms			

	75% Percentile: 1.03 +/- 0.01 ms 5% Percentile: 0.88 +/- 0.01 ms 95% Percentile: 1.46 +/- 0.08 ms **60GHz-C0U0-1400** Mean: 1.49 +/- 0.29 ms Standard deviation: 3.85 +/- 2.62 ms Median: 1.05 +/- 0.01 ms Min: 0.89 +/- 0.02 ms Max: 41.25 +/- 27.14 ms 25% Percentile: 1.00 +/- 0.01 ms 75% Percentile: 1.11 +/- 0.01 ms 5% Percentile: 0.94 +/- 0.01 ms 95% Percentile: 1.60 +/- 0.13 ms
Complementary measurement results	n/a

A.5.4.2 Throughput

Annex Table A-62 Throughput between nodes interconnected via 60GHz Metrolinq System

Test Case ID	TC_THR_Tcp / TC_THR_Udp			
General description of the test	The test assesses the average Throughput between two communication end-points (VNF) deployed on both sides of a 60GHz backhaul.			
Purpose	The test acts as a calibration test to primarily assess the performance of the intermediate 60GHz backhaul link.			
Executed by	Partner:	FOKUS	Date:	09.03.2020
Involved Partner(s)	FOKUS			
Scenario	The communication endpoints of the service (VNFs), i.e. the virtual instruments, are placed within the FOKUS02 and NetApp HCI, both interconnected via a Dell Switch and the 60 GHz system. The test is repeated for the up- and down-link as well as for having one vs. four parallel data streams			
Slicing configuration	An exclusive slice / software-defined network is used for the test.			
Components involved	• ThinkCenter with probe • Dell Switch • NetApp HCI • The connecting 60GHz backhaul; MetroLinq nodes.			
Metric(s) under study	*Throughput*			
Additional tools involved	TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints			
Primary measurement results	**Throughput** **C0U0-60GHz-TC_THR_Tcp** Mean: 901.02 +/- 9.25 Mbps Standard deviation: 13.77 +/- 1.78 Mbps			

	Median: 901.86+/- 9.60 Mbps Min: 819.24 +/- 15.07 Mbps Max: 926.36 +/- 5.94 Mbps 25% Percentile: 893.77 +/- 10.80 Mbps 75% Percentile: 910.59 +/- 7.38 Mbps 5% Percentile: 881.19 +/- 12.39 Mbps 95% Percentile: 918.49 +/- 6.50 Mbps **C0U0-60GHz-TC_THR_Udp** Mean: 901.00 +/- 9.08 Mbps Standard deviation: 12.33 +/- 1.76 Mbps Median: 901.38 +/- 9.39 Mbps Min: 857.24 +/- 11.71 Mbps Max: 937.48 +/- 7.94 Mbps 25% Percentile: 893.10 +/- 10.66 Mbps 75% Percentile: 910.17 +/- 7.30 Mbps 5% Percentile: 881.12 +/- 12.03 Mbps 95% Percentile: 918.34 +/- 6.54 Mbps **60GHz-C0U0-TC_THR_Tcp** Mean: 907.18 +/- 3.99 Mbps Standard deviation: 18.44 +/- 1.59 Mbps Median: 909.94 +/- 3.81 Mbps Min: 819.00 +/- 13.00 Mbps Max: 941.68 +/- 1.30 Mbps 25% Percentile: 897.96 +/- 5.08 Mbps 75% Percentile: 919.86 +/- 2.88 Mbps 5% Percentile: 875.43 +/- 7.09 Mbps 95% Percentile: 930.70 +/- 1.68 Mbps **60GHz-C0U0-TC_THR_Udp** Mean: 906.62 +/- 1.56 Mbps Standard deviation: 24.85 +/- 4.81 Mbps Median: 908.60 +/- 1.81 Mbps Min: 749.80 +/- 88.11 Mbps Max: 969.76 +/- 5.80 Mbps 25% Percentile: 895.20 +/- 2.15 Mbps 75% Percentile: 920.64 +/- 1.24 Mbps 5% Percentile: 873.26 +/- 2.73 Mbps 95% Percentile: 935.46 +/- 1.31 Mbps

A.5.5 60 GHz IHP Prototype

A.5.5.1 Round Trip Time

Annex Table A-63 RTT between nodes interconnected via 60GHz IHP System

Test Case ID	TC_RTT_e2e
General description of the test	The test assesses the average, minimum, and maximum RTT between two communication endpoints (VNF) deployed on both sides of a 60 GHz backhaul.

<table>
<tr><td>Purpose</td><td colspan="4">The test acts as a calibration test to primarily assess the performance of the intermediate 60GHz backhaul link.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>FOKUS</td><td>Date:</td><td>06.03.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">FOKUS, IHP</td></tr>
<tr><td>Scenario</td><td colspan="4">The communication endpoints of the service (VNFs), i.e. the virtual instruments, are placed on two machines interconnected by the IHP 60 GHz backhaul and a local switch</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">An exclusive slice / software-defined network is used for the test.</td></tr>
<tr><td>Components involved</td><td colspan="4">• 2 local servers acting as communication endpoints
• Dell Switch
• The connecting 60 GHz backhaul is realized by the IHP nodes.</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4"><i>RTT</i></td></tr>
<tr><td>Additional tools involved</td><td colspan="4">TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication end-points</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">Round Trip Time
IHP60GHz_src-IHP01_dst-IHP03-32
Mean: 0.31 +/- 0.01 ms
Standard deviation: 0.05 +/- 0.00 ms
Median: 0.29 +/- 0.00 ms
Min: 0.250 +/- 0.00 ms
Max: 0.59 +/- 0.04 ms
25% Percentile: 0.27 +/- 0.00 ms
75% Percentile: 0.32 +/- 0.01 ms
5% Percentile: 0.26 +/- 0.00 ms
95% Percentile: 0.40 +/- 0.01 ms
IHP60GHz_src-IHP01_dst-IHP03-56
Mean: 0.31 +/- 0.01 ms
Standard deviation: 0.05 +/- 0.00 ms
Median: 0.30 +/- 0.01 ms
Min: 0.25 +/- 0.00 ms
Max: 0.57 +/- 0.05 ms
25% Percentile: 0.28 +/- 0.00 ms
75% Percentile: 0.33 +/- 0.02 ms
5% Percentile: 0.26 +/- 0.00 ms
95% Percentile: 0.41 +/- 0.02 ms
IHP60GHz_src-IHP01_dst-IHP03-1400
Mean: 0.83 +/- 0.043 ms
Standard deviation: 0.043 +/- 0.00 ms
Median: 0.83 +/- 0.04 ms
Min: 0.77 +/- 0.043 ms
Max: 1.04 +/- 0.05 ms
25% Percentile: 0.81 +/- 0.04 ms
75% Percentile: 0.84 +/- 0.047 ms
5% Percentile: 0.78 +/- 0.04 ms
95% Percentile: 0.91 +/- 0.04 ms</td></tr>
<tr><td>Complementary measurement results</td><td colspan="4">n/a</td></tr>
</table>

A.5.5.2 Throughput

Annex Table A-64 Throughput between nodes interconnected via 60 GHz IHP System (Uplink IHP01-IHP03; Downlink IHP03-iHP01)

<table>
<tr><td>Test Case ID</td><td colspan="4">TC-_THR_Tcp TC_THR_Udp</td></tr>
<tr><td>General description of the test</td><td colspan="4">The test assesses the average Throughput between two communication endpoints (VNF) deployed on both sides of a 60GHz backhaul.</td></tr>
<tr><td>Purpose</td><td colspan="4">The test acts as a calibration test to primarily assess the performance of the intermediate 60GHz backhaul link.</td></tr>
<tr><td>Executed by</td><td>Partner:</td><td>FOKUS</td><td>Date:</td><td>09.03.2020</td></tr>
<tr><td>Involved Partner(s)</td><td colspan="4">FOKUS</td></tr>
<tr><td>Scenario</td><td colspan="4">The communication endpoints of the service (VNFs), i.e. the virtual instruments, are placed within the IHP-01 and IHP-03, both interconnected via a Dell Switch and the 60 GHz system.
The test is repeated for the up- and down-link as well as for having one vs. four parallel data streams</td></tr>
<tr><td>Slicing configuration</td><td colspan="4">An exclusive slice / software-defined network is used for the test.</td></tr>
<tr><td>Components involved</td><td colspan="4">• 2 local servers acting as measurement endpoints
• Dell Switch
• The connecting 60GHz backhaul ; IHP nodes.</td></tr>
<tr><td>Metric(s) under study</td><td colspan="4"><i>Throughput</i></td></tr>
<tr><td>Additional tools involved</td><td colspan="4">TAP for automated testing, Debian-based virtual instruments (VNFs) acting as communication endpoints</td></tr>
<tr><td>Primary measurement results</td><td colspan="4">Throughput
IHP60GHz_src-IHP01_dst-IHP03-TC-Thr-Tcp
Mean: 864.90 +/- 0.03 Mbps
Standard deviation: 0.51 +/- 0.03 Mbps
Median: 865.00 +/- 0.00 Mbps
Min: 863.76 +/- 0.22 Mbps
Max: 866.00 +/- 0.00 Mbps
25% Percentile: 864.95 +/- 0.08 Mbps
75% Percentile: 865.0 +/- 0.00 Mbps
5% Percentile: 863.99 +/- 0.02 Mbps
95% Percentile: 865.63 +/- 0.19 Mbps
IHP60GHz_src-IHP01_dst-IHP03-TC-Thr-Udp
Mean: 863.36 +/- 0.02 Mbps
Standard deviation: 0.63 +/- 0.03 Mbps
Median: 863.00 +/- 0.00 Mbps
Min: 862.60 +/- 0.21 Mbps
Max: 865.72 +/- 0.22 Mbps
25% Percentile: 863.0 +/- 0.00 Mbps
75% Percentile: 863.92 +/- 0.11 Mbps
5% Percentile: 862.97 +/- 0.03 Mbps
95% Percentile: 864.0 +/- 0.00 Mbps</td></tr>
</table>

	IHP60GHz_src-IHP03_dst-IHP01-TC-Thr-Tcp Mean: 892.26 +/- 0.03 Mbps Standard deviation: 0.69 +/- 0.11 Mbps Median: 892.00 +/- 0.00 Mbps Min: 890.48 +/- 0.82 Mbps Max: 893.04 +/- 0.08 Mbps 25% Percentile: 892.00 +/- 0.00 Mbps 75% Percentile: 893.00 +/- 0.00 Mbps 5% Percentile: 891.30 +/- 0.17 Mbps 95% Percentile: 893.00 +/- 0.00 Mbps **IHP60GHz_src-IHP03_dst-IHP01-TC-Thr-Udp** Mean: 892.46 +/- 0.02 Mbps Standard deviation: 0.99 +/- 0.03 Mbps Median: 892.20 +/- 0.15 Mbps Min: 891.00 +/- 0.00 Mbps Max: 896.60 +/- 0.29 Mbps 25% Percentile: 892.00 +/- 0.00 Mbps 75% Percentile: 893.00 +/- 0.00 Mbps 5% Percentile: 891.16 +/- 0.14 Mbps 95% Percentile: 893.25 +/- 0.15 Mbps
Complementary measurement results	n/a

A.5.5.3 Smart Grid Control Traffic

Annex Table A-65 GOOSE Control message latency – KVM scenario

Test Case ID	TC_LatSmartGridControlMsgLatency_BER			
General description of the test	The tests evaluate the latency GOOSE messages experience when being sent over a 5G Core.			
Purpose	Determine the latencies GOOSE messages experience over a 5G Core, and how these latencies are impacted by the deployment model used and by the transmission-interval lengths.			
Executed by	Partner:	KAU	Date:	27.05.2020
Involved Partner(s)	KAU, FhG			
Scenario	KVM-based deployment			
Slicing configuration	N/A			
Components involved	• 2 GRE-enabled GOOSE gateways • 1 machine running the Open5GCore platform • - 1 machine running the GOOSE emulation software			
Metric(s) under study	*Latency* *(The target KPI Latency is Average E2E GOOSE Message Latency and the Average 5GCore Latency, as defined in TC-Lat-X)*			
Additional tools involved	N/A			

Primary measurement results	Average E2E GOOSE Message Latency [ms],			
		Transmission interval of 10 ms	Transmission interval of 20 ms	Transmission interval of 50 ms
	Mean	0.84 ± 0.01	0.85 ± 0.02	0.85 ± 0.03
Complementary measurement results	Average 5GCore Latency [ms],			
		Transmission interval of 10 ms	Transmission interval of 20 ms	Transmission interval of 50 ms
	Mean	0.49 ± 0.02	0.49 ± 0.02	0.51 ± 0.02

Annex Table A-66 GOOSE Control message latency – Docker scenario

Test Case ID	TC_LatSmartGridControlMsgLatency_BER			
General description of the test	The tests evaluate the latency GOOSE messages experience when being sent over a 5G Core.			
Purpose	Determine the latencies GOOSE messages experience over a 5G Core, and how these latencies are impacted by the deployment model used and by the transmission-interval lengths.			
Executed by	Partner:	KAU	Date:	27.05.2020
Involved Partner(s)	KAU, FhG			
Scenario	Docker-based deployment			
Slicing configuration	N/A			
Components involved	• 2 GRE-enabled GOOSE gateways • 1 machine running the Open5GCore platform • 1 machine running the GOOSE emulation software			
Metric(s) under study	*Latency* *(The target KPI Latency is Average E2E GOOSE Message Latency and the Average 5GCore Latency, as defined in TC-Lat-XXX)*			
Additional tools involved	N/A			
Primary measurement results	Average E2E GOOSE Message Latency [ms],			
		Transmission interval of 10 ms	Transmission interval of 20 ms	Transmission interval of 50 ms
	Mean	0.56 ± 0.01	0.55 ± 0.02	0.57 ± 0.01
Complementary measurement results	Average 5GCore Latency [ms],			
		Transmission interval of 10 ms	Transmission interval of 20 ms	Transmission interval of 50 ms
	Mean	0.23 ± 0.01	0.23 ± 0.01	0.23 ± 0.01