

A new approach to 5G and MEC integration

Lechośław Tomaszewski*, Sławomir Kukliński*[◇], Robert Kołakowski*
 *Orange Polska, [◇]Warsaw University of Technology

Abstract

In the paper a concept of integration of MEC into the 5G network slicing architecture has been presented. Three variants of the architecture have been proposed, which incorporate: individual MEP/MEPM for each slice, shared ones for multiple network slices and Distributed Autonomous Slice Management and Orchestration (DASMO) approach. Each variant is focused on efficient integration of 5G Core and MEC solutions and utilizing additional functionalities of both components, including MEC APIs and 5G Control Plane exposure capabilities. Finally, main issues of 5G-MEC implementation have been discussed, which include the aspects of MEC service APIs, MEC Apps mobility in demanding use cases, challenges of service continuity in roaming scenarios, as well as the role and availability of 5G enablers.

Index Terms

5G, MEC, Network slicing, Orchestration, Management, Architecture, Scalability

I. INTRODUCTION

ONE of the main challenges, since the introduction of 5G network, remains meeting the requirements of latency-critical services. Initially, network slicing together with RAN enhancements were perceived as satisfactory enablers for achieving much lower delay than previous generation networks. However, due to the cloud character of networks and related time consumed by physical transmission of data, it has become clear that in order to meet the exorbitant requirements for low latency, services have to be migrated to the edge of the network.

Currently, the integration of ETSI Multi-access Edge Computing (MEC) and 5G is considered as a possible solution to the problem. Despite considerable facilitation provided by MEC, such as application deployment close to UE or mobility procedures enabling migration of applications to other hosts for maintaining low latency, the current status of MEC concept development raises significant concerns.

In the paper issues of MEC in terms of integration with slice-enabled, stand-alone 5G System (5GS) have been discussed. Such an integration has already been described by ETSI; however, it seems that this approach to integration of MEC with NFV and slicing-enabled 5G network is over-complex. It is hereby proposed to use a different way, by removing duplicate functionalities of both solutions and integrating the information provided by network slice instances (NSIs) control message bus and MEC APIs. It is also proposed to integrate the approach with In-Slice Management (ISM) concept. The paper concerns not solely the architectural alterations but also a discussion on mechanisms of MEC and 5G for their reciprocal integration.

The structure of the paper is as follows. Section 2 describes the current state of the art of MEC and its facilitation towards integration with 5G network. In Section 3 the scalable MEC-enabled network slicing architecture is proposed. In Section 4 implementation remarks are presented. Section 5 summarizes and concludes the paper.

II. STATE OF THE ART

Recently, an intensified research can be observed in the field of integration of the MEC framework with the 5G Core (5GC). Basic requirements concerning MEC–5GC inter-operation have been defined in [1]. The most important ones involve necessary MEC support for provisioning of traffic steering and policy control information of applications in 5GS. Specific data have to be exchanged between MEC management entities and 5G Network Exposure Function (NEF).

Following the basic requirements, a common architecture integrating MEC and NFV has been defined [2], which incorporates ETSI NFV MANO as a part of the management and orchestration domain. To prevent doubling of management functionalities, vital changes have been introduced, which include i.a. moving Life Cycle Management (LCM) of the MEC platform and MEC Apps to NFVO and VFNM. Moreover, the NFV Infrastructure (NFVI) is used for deployment of MEC Apps, MEC Platforms and MEC Platform Managers.

Two essential enhancements for integration of MEC with 5G network slicing have been introduced in [3]. According to the first one, a slice is deployed in the MEC ecosystem, meaning that MEC Platform (MEP) and MEC Platform Manager (MEPM) entities can be shared by several slices. The second one involves MEC deployment within a slice – each slice has its separate MEP and MEPM entities. Deployment of multiple slices within MEC raises fundamental questions concerning performance, isolation and slice awareness. According to [3], to enable safe operation, slice awareness of MEC Applications Orchestrator

(MEAO) and virtualized MEPM (MEPM-V) as well as isolation of each slice (in case of shared MEP/MEPM approach) have to be ensured. Separation has to be provided especially in terms of MEC Apps access to UEs information.

Another important aspect of 5G-MEC integration is provisioning of mechanisms for seamless migration of applications and thus service continuity in latency-critical scenarios. Exemplary solutions have been proposed in [5], such as MEC host pre-allocation based on UE trajectory prediction or creating *relocation groups* – sets of MEC hosts in UE’s vicinity where each host is pre-configured for running desired application to reduce deployment time in case of handover.

Ksentini *et al.* have proposed the architecture of MEC-enabled 5G network [4] where MEC is defined as a separate orchestration domain but implemented on the same NFVI as VNFs and hosted by the ETSI NFV MANO stack of the VNF domain. Several aspects concerning security and data isolation in MEC APIs have been pointed out, i.a. filtering of RAN-related data before delivery to MEP in in-slice deployment scenarios. The results of exemplary MEC Apps deployment times in a laboratory environment have been presented.

An interesting view of the reasons for MEC implementation and its advantages is proposed in [6]. The additional benefits to usually emphasized latency reduction and UE’s computational offloading, include data scalability, application scalability, intent-driven networking or partial offloading of network functions. They pertain directly to optimization of network architectures, distribution of traffic (especially in the transmission domain), network operations and – in result – both investment and operational costs, which are issues of premium importance for network operators.

III. SCALABLE MEC-ENABLED NETWORK SLICING ARCHITECTURE

A. Fundamentals of the concept

Kukliński *et al.* have proposed a reference architectural framework for network slicing [7], which is based on the ETSI NFV MANO architecture [8], compliant with various communication network architectures and facilitates vertical and horizontal slice expansion due to incorporation of common/dedicated slice concepts, exposure of slice functions via slice API and slice stitching. The framework follows the paradigm of hierarchical multi-domain orchestration and supports tenant-oriented operations and interfaces based on embedded in-slice managers. In [9] the internal structure of slices has been further defined – the core part of slice, consisting of functions composing the Application (AP), Control (CP) and Data (DP) Planes (A-VNFs, C-VNFs and S-VNFs, respectively), is accompanied by two special functional blocks: Slice Manager (SM) and Slice Operation Support (SOS), both implemented as sets of VNFs (M-VNFs and S-VNFs respectively), belonging to slice template and sharing the lifecycle of their slice. The architecture is called “Distributed Autonomous Slice Management and Orchestration” (DASMO) and is presented in Fig. 1.

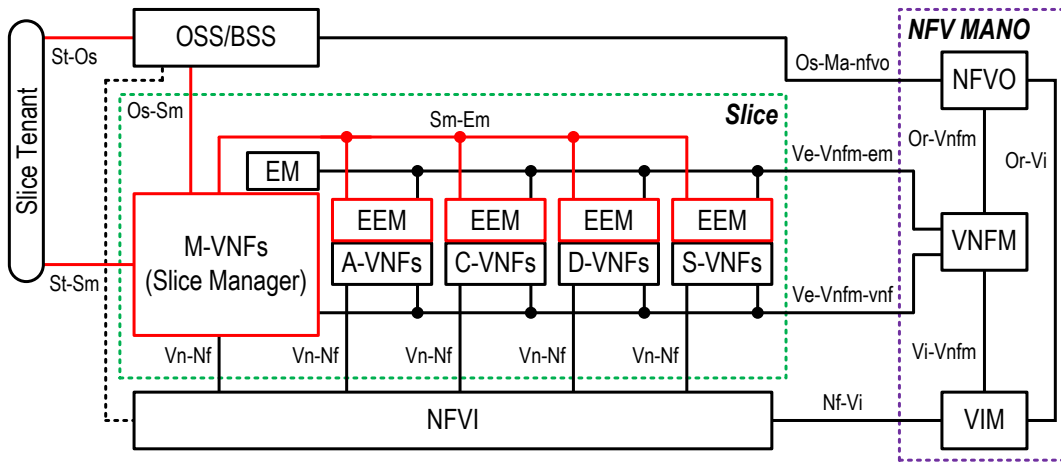


Fig. 1. DASMO framework with internal structure of network slices – ETSI NFV MANO extensions (slice management plane shown in red)

SM is a central point of slice management plane and has links to Embedded Element Managers (EEMs) of all VNFs implemented within a slice. These EEMs follow the ETSI NFV concept of Element Manager (EM), but they are augmented with additional functionalities facilitating slice-level management support, VNF monitoring, actuating and autonomic control loop, etc. SM plays a role of slice OSS and incorporates the functions responsible for slice-level monitoring, analysis, actuating and autonomic control loop according to the Monitor-Analyze-Plan-Execute (MAPE) [10] model (real-time feedback loop). Additionally, SM implements tenant-oriented functions: accounting, KPI monitoring and reporting, configuration support (following the “intent-based management” paradigm), which are exposed via the Tenant Portal functionality of SM. SM exposes also an interface to the global OSS/BSS, which is of importance especially in multi-domain slicing. SOS functions support slice-level operations as slice selection, subscription, authentication and stitching of sub-slices to provide transparent communication between NFs belonging to different domains for creation of the end-to-end slice.

The described architecture implements the ISM concept, which – due to hierarchical distribution of management tasks – is inherently scalable. The scalability of orchestration may be achieved by using recursive orchestration (“MANO in MANO”), and the DASMO concept is compliant with it.

B. Inclusion of the MEC framework into the 5G network slicing architecture

The proposed MEC-enabled 5G network slicing architecture is based on the following principles:

- MEC services, similarly as NSIs, have limited geographic scope and are focused on a specific service – this is in line with the network slicing philosophy, which emphasizes customization of NSI to its service or a group of services with similar characteristics. In more complicated use cases, like UAV or V2X, the overall service uses several NSIs of different type. Utilization of MEC as platform offers useful mechanisms to provide a specific service. Consequently, in case of network slicing, the number of MEC Apps will be limited and they will be defined during the slice creation. Therefore, orchestration of MEC Apps during the NSI run-time will be rather rare.
- Flexible architectural approach, adapted to NSI characteristics (complexity, longevity, critical deployment time, etc.), is required. As a result, coexistence of various architectural variants can be expected.
- Implementation of MEC Apps as a part of slice AP – the same NFVI is used by CP/DP and no separated MEC orchestration domain is needed. Therefore, the orchestration of MEC Apps belongs to slice-level orchestration activities.
- Tight integration on an equal basis of MEC APIs (RNIS, Localization, etc.) with information obtainable from 5GC via NEF, to extend the amount of information available for slice creation and for avoidance of duplication of 5G and MEC functions like Network Repository Function (NRF), etc.

Fig. 2 shows the proposed generalized architecture of MEC and 5G integration. All VNFs are implemented in the VNF space, using common NFVI managed by VIM (omitted in the picture for simplification). NFVI can be single- or multi-domain (cf. [11]). All VNFs have their EMs (symbolized by red dots) connected to OSS/BSS (red arrows). In case of MEC Apps, their management functions may be embedded in applications, externalized or nonexistent. VNFs and their EMs are also connected to VNFM(s) (single- or multi-VNFM options are possible, cf. [11]), which are responsible for LCM of both MEC Apps and other VNFs (VNFM* in Fig. 2). Even if the ETSI MEC framework assumes *Ve-Vnfm-vnf* variant (light) of MEC App–NFVO reference point, it may be potentially useful in specific cases to implement fully functional *Ve-Vnfm-em* variant, instead.

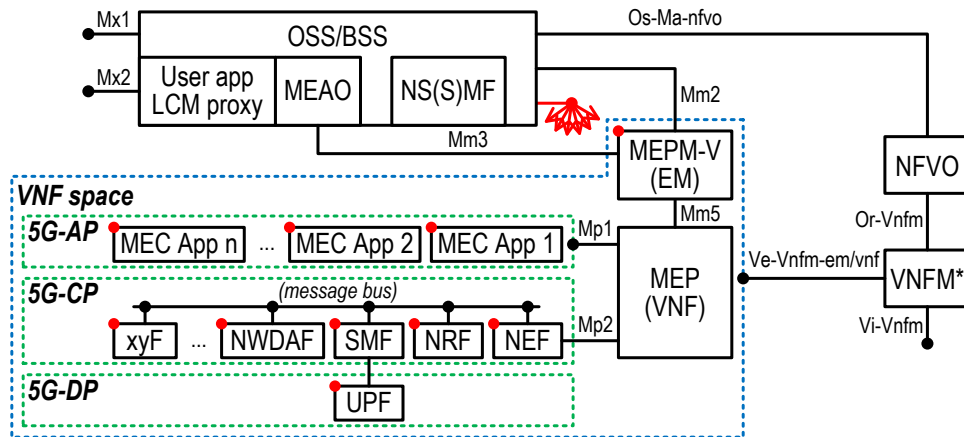


Fig. 2. General slicing architecture of MEC-enabled 5G network

Orchestration of MEC is located at OSS/BSS together with management of a 5G network and Network Slice (Subnet) Management Function – NS(S)MF. Therefore, all interactions with the ETSI NFV MANO stack are performed via one common OSS–NFVO interface. As MEAO and User app LCM proxy are functional modules of OSS/BSS, some ETSI MEC reference points are internalized. OSS/BSS opens both interfaces *Mx1/Mx2* to the customer domain. MEP exposes platform’s services to MEC Apps (*Mp1*) and in case of 5GS-interacting ones, acts as mediator to 5GC-CP via NEF (*Mp2*, considered as *Naf* at the 5GC-CP bus).

The described generalized architecture is valid both in case of 5G network with its own MEP/MEPM-V (Variant 1) and for MEP/MEPM-V sharing by multiple networks (Variant 2). In case of Variant 1, the “VNF space” in Fig. 2 can be simply renamed to “5G network”. As MEP/MEPM-V are dedicated, they can be a part of the template of virtualized 5G network and share its life cycle. In case of Variant 2 (suitable rather for short-lived and simple slices), they will be external to 5G networks (now consisted of AP, CP and DP only). As the shared MEP is interfaced with CPs and APs of separate networks, it has to provide mechanisms for mutual isolation between these networks, i.e. their reciprocal unawareness and prevention of cross-exchange of information or unauthorized access to foreign 5GC-CP. The issue of protection of individual networks privacy is an additional factor for externalization of MEP towards all connected networks in Variant 2. Additionally, inter-App privacy

should be ensured in both variants (e.g. awareness of users, their sessions metadata, etc.), but it can be provided by the own 5GC-CP. If network slicing is enabled (the case of multiple-NSI networks, providing services with different characteristics), both MEP/MEPM-V and MEAO have to be NSI-aware, i.e. recognize and distinguish NSIs, as it is required from all 5GC-CP entities (cf. [3]).

C. Scalable MEC-enabled slicing architecture

In geographically distributed architecturally complex communication networks, moving network functions of high granularity towards the edge has positive consequences for user traffic transport and performance but at the expense of the control and management planes. Centralized management of highly distributed networks is vastly inefficient, especially due to necessity of transporting of huge volumes of data needed for analysis, decision-making and execution of automated management processes. The DASMO architecture faces this problem.

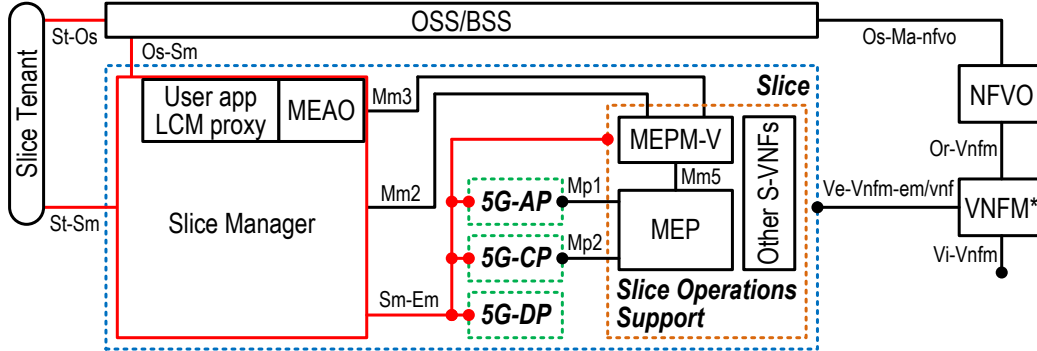


Fig. 3. MEC-enabled DASMO architecture

The single-domain scalable MEC-enabled slicing architecture (DASMO extended with MEC) is presented in Fig. 3. All VNFs of the slice have their own EEMs, as it is required by the DASMO concept. EEMs are connected to SM, to provide the slice management plane communication. MEP/MEPM-V belong to the SOS area, because their role is in line with the SOS definition, especially the exposure of transparent mechanisms for slice VNFs interconnection. MEAO and User app LCM proxy are located in SM, because it plays a role of slice OSS.

The significant task of SM is proper routing of the MEC framework-related exchange. The $Mm1$ communication will be forwarded to the global OSS/BSS, which concentrates the exchange with NFV MANO. The $Mx1/Mx2$ reference point communication will be exposed through the $St-Sm$ interface. Alternatively, it may be forwarded to the global OSS/BSS if the Slice Tenant prefers interactions that way (e.g. utilization of multiple separate NSIs; the consolidated global view is then desired).

It has to be noted that the DASMO architecture also supports the multi-domain sliced networks. The global OSS/BSS contains the Multi-Domain Management and Orchestration Support functions composed of Multi-Domain Slice Configurator (MDSC) and Multi-Domain Orchestrator (“Umbrella NFVO”, cf. [11]). MDSC, during the slice run-time, keeps monitoring of the end-to-end slice and coordinates its reconfiguration, taking also care of MEC-related activities. It is responsible for proper configuration of local SOS entities for inter-domain operations.

To enable operations in a multi-domain environment, it is essential to provide means of horizontal end-to-end slice stitching, i.e. concatenation of sub-slices from different domains. Inter-Domain Operations Support (IDOS), a functional part of SOS, is defined for this purpose. IDOS acts as an inter-slice gateway, implementing information exchange between neighbouring domains, i.e. exposure of domain abstracted view and support for inter-domain communication (relevant protocols, transcoding, mediation, etc.). In the MEC-enabled DASMO architecture, the $Mp3$ reference point control information transfer between MEPs shall be carried out via IDOS.

IV. 5G-MEC IMPLEMENTATION DETAILS

A. MEC service APIs

The MEC framework defines special service APIs exposed by MEP to MEC Apps: Radio Network Information – RNIS [12] (PLMN information, E-RAB information, S1 Bearer information and L2 measurements), Location [13] (zonal presence and terminal location, including information about distance from specific location or between terminals), UE Identity [14] and Bandwidth Management [15] (management of bandwidth on per application session basis). These services shall be provided via the $Mp2$ reference point, which will need special enablers within 5GC-CP. It has to be noted that the ETSI MEC framework is currently defined for integration with the 4G network (it is especially reflected in RNIS data model, which is not radio technology-agnostic). Therefore, specifications of these APIs have to be updated and corresponding 5GS-side enablers have to

be available. This mainly applies to mechanisms provided by NEF, NWDAF [16] and LCS [17]. It is particularly important to ensure the availability of RAN-related information. Although the 5G RAN physical layer measurements at UE have been specified [18], the mechanisms similar to 3G/4G radio measurements collection (MDT, cf. [19]) for further processing and use are still undefined, but they are in the scope of Release 17.

Additionally, it is hereby proposed to define the special MEP-facing gateway function located in 5GC-CP to provide a single and standardized interface for MEP and ensure smooth and optimized interaction (especially for avoiding excessive signaling exchange within 5GC-CP). Such initiative needs bilateral cooperation of the 3GPP and ETSI MEC group.

B. Application mobility in demanding use cases

Ksentini *et al.* demonstrated [4] that the total time needed for MEC application deployment can vary from ~ 60 s (application instantiation only) to ~ 180 s (onboarding and instantiation) or even ~ 440 s (full onboarding and instantiation of both MEP and application). In high-mobility use cases (drone, railway or automotive – speeds of several kms per minute) MEC Apps cannot only follow the UE, but they must overtake it. Standard mechanisms of location tracking (even with prediction) will not be sufficient, so integration with drone traffic management system (aware of flight plan), with UE context awareness mechanisms (driven by mechanisms of Artificial Intelligence and Geographic Information Systems to deduce e.g. following a motorway or railway line) or with on-board navigation (aware of the desired route) can be utilized.

C. Service continuity in roaming

Special concern should be dedicated to roaming cases. Maintaining service continuity requires replication of its architecture at VPLMN and new approach to the re-registration process when changing an operator. This issue is partially discussed in [20]. In case of MEC-enabled service architecture, the entire NSI, along with the MEC App residing in the AP, must be instantiated on VPLMN resources in local breakout mode. To some extent, service architectures (i.e. NSI templates) standardization together with MEC applications porting mechanism can be a solution, but a general mechanism for any NSI portability will be required.

D. Availability of 5G enablers for MEC

Majority of R&D projects are based on popular 5GS implementations, such as OpenAirInterface, Open5GCore or free5GC. It has to be noted that these solutions implement fundamental functionalities of the 3GPP 5G architecture, but unfortunately NEF, NWDAF or LCS are missing there. Even handover support can be somewhat problematic. Additionally, whenever non UE-based positioning is required, the Network-Assisted Positioning Procedure shall be used, which has to be supported by gNB (positioning based on RAN measurements, cf. [21]). Individual efforts on implementation of these mechanisms or an initiative on public-domain tools are needed. The list, review and status of open source tools for 5G (3GPP Release 15) can be found in [22].

V. SUMMARY AND CONCLUSIONS

In this paper, a novel concept of scalable MEC integration with slicing-enabled 5G has been presented. It is based on the critical evaluation of the ETSI approach. Several issues of the concept of ETSI, such as separate orchestration of network slices and MEC or improper integration of services offered by the 5GC-CP bus and MEC APIs, have been identified. Moreover, there is a need of filtering of MEC data on per slice basis to provide proper isolation between slices.

It has been proposed to use an ETSI NFV MANO-compliant orchestration based on a single orchestrator, common both for MEC and NFV, and to follow the ISM paradigm for automation of MEC operations and providing slice management scalability. Also, the specifics of use of MEC for network slicing has been outlined – the region of operation is already defined by the slice registration area. Moreover, slice services are generally limited and they rarely change during slice lifetime. The observations have led to the conclusions that it is viable to use many simplifications of MEC and slicing-enabled 5G network integration. The main value of MEC in the context of network slicing is the set of APIs, especially RNIS, Localization and Application Mobility API. Further research and development directions for enabling of commercial implementation and use of the presented concept have been also outlined.

The three defined and described variants of integration of MEC and 5G network may coexist in the network. Their selection is dependent on slice duration, complexity, deployment time requirements or the expected number of slices. The ISM-based integration is recommended as a variant of choice.

REFERENCES

- [1] ETSI: ETSI GS MEC 002 Multi-access Edge Computing (MEC); Phase 2: Use Cases and Requirements, V2.1.1. ETSI MEC ISG (Oct 2018)
- [2] ETSI: ETSI GS MEC 003 Multi-access Edge Computing (MEC); Framework and Reference Architecture, V2.1.1. ETSI MEC ISG (Jan 2019)
- [3] ETSI: ETSI GR MEC 0024 Multi-access Edge Computing (MEC); Support for network slicing, V2.1.1. ETSI MEC ISG (Nov 2019)
- [4] Ksentini A., Frangoudis P. A.: Towards Slicing-Enabled Multi-access Edge Computing in 5G. *IEEE Network* **34**(2), 99–105
- [5] ETSI: ETSI GR MEC 018 Multi-access Edge Computing (MEC); End to End Mobility Aspects, V1.1.1. ETSI MEC ISG (Oct 2017)(2020)

- [6] Iwai T., Nakao A.: Demystifying myths of MEC: rethinking and exploring benefits of Multi-access/Mobile Edge Computing. In 2018 IEEE 7th International Conference on Cloud Networking (CloudNet), pp. 1–4. IEEE, Tokyo (2018)
- [7] Kukliński S., Tomaszewski L. *et al.*: A reference architecture for network slicing. In: 2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft), pp. 217–221. IEEE, Montreal (2018).
- [8] ETSI: ETSI GS NFV 002 Network Functions Virtualisation (NFV); Architectural Framework, V1.2.1. ETSI NFV ISG (Dec 2014)
- [9] Kukliński S., Tomaszewski L.: DASMO: A scalable approach to network slices management and orchestration. In: NOMS 2018 – 2018 IEEE/IFIP Network Operations and Management Symposium, pp. 1–6. IEEE, Taipei (2018).
- [10] IBM: An architectural blueprint for autonomic computing. Autonomic Computing White Paper, Third Edition (Jun 2005)
- [11] ETSI: ETSI GS NFV-IFA 009 Network Functions Virtualisation (NFV); Management and Orchestration; Report on Architectural Options, V1.1.1. ETSI NFV ISG (Jul 2016)
- [12] ETSI: ETSI GS MEC 012 Multi-access Edge Computing (MEC); Radio Network Information API, V2.1.1. ETSI MEC ISG (Dec 2019)
- [13] ETSI: ETSI GS MEC 013 Multi-access Edge Computing (MEC); Location API, V2.1.1. ETSI MEC ISG (Sep 2019)
- [14] ETSI: ETSI GS MEC 014 Mobile Edge Computing (MEC); UE Identity API, V1.1.1. ETSI MEC ISG (Feb 2018)
- [15] ETSI: ETSI GS MEC 015 Mobile Edge Computing (MEC); Bandwidth Management API, V1.1.1. ETSI MEC ISG (Oct 2017)
- [16] 3GPP: 3GPP TS 23.501 System Architecture for the 5G System, V16.4.0, 3GPP (Mar 2020)
- [17] 3GPP: 3GPP TS 23.273 5G System (5GS) Location Services (LCS); Stage 2, V16.3.0. 3GPP (Mar 2020)
- [18] 3GPP, 3GPP TS 38.215 NR; Physical layer measurements, V16.1.0. 3GPP (Apr 2020)
- [19] 3GPP, 3GPP TS 37.320 Universal Terrestrial Radio Access (UTRA) and Evolved Universal Terrestrial Radio Access (E-UTRA); Radio measurement collection for Minimization of Drive Tests (MDT); Overall description; Stage 2, V15.0.0. 3GPP (Jul 2018)
- [20] Tomaszewski L., Kolański R., Korzec P.: On 5G support of cross-border UAV operations. In: 2020 IEEE International Conference on Communications, Workshop on Integrating UAVs into 5G and Beyond (accepted)
- [21] 3GPP: 3GPP TS 38.305 NG Radio Access Network (NG-RAN); Stage 2 functional specification of User Equipment (UE) positioning in NG-RAN, V15.5.0. 3GPP (Jan 2020)
- [22] 5G Americas: The Status of Open Source for 5G. 5G Americas White Paper (Feb 2019)