EURECOM
Department of Mobile Communications
Campus SophiaTech
CS 50193
06904 Sophia Antipolis cedex
FRANCE

Research Report RR-16-312

Analyzing MEC Architectural Implications for LTE/LTE-A

February 2th, 2016
Last update May 4th, 2016

Chia-Yu Chang, Konstantinos Alexandris, Navid Nikaein, and Thrasyvoulos Spyropoulos

Tel : (+33) 4 93 00 81 00
Fax : (+33) 4 93 00 82 00
Email : {Firstname.Name}@eurecom.fr

[1]EURECOM's research is partially supported by its industrial members: BMW Group Research and Technology, IABG, Monaco Telecom, Orange, Principaut de Monaco, SAP, SFR, ST Microelectronics, Symantec.

Analyzing MEC Architectural Implications for LTE/LTE-A

Chia-Yu Chang, Konstantinos Alexandris, Navid Nikaein, and Thrasyvoulos Spyropoulos

Abstract

Lower-latency services are expected toward beyond 4G mobile networks; however, legacy 3G and 4G networks are now suffering from the mobile data surge. In this sense, pushing network services to the network edge has the potential to improve the traffic latency and user experience, as well as to offload Internet traffic. The network functions can be dynamically re-assigned between edge and core nodes, offering flexibility and scalability. Even such Mobile Edge Computing (MEC) principles can highly benefit to the LTE/LTE-A network, but a detailed MEC architecture is not currently in place. To this end, in this research report, we analyze the impact of MEC on the current LTE/LTE-A systems, considering use-cases of interest, identied by ETSI. Finally, a modular MEC architecture is proposed, identifying the required radio network interfaces, application development services, and application programming interface.

Index Terms

Mobile-edge computing, 4G, 5G, cloud, low-latency service.

Contents

1 Introduction **1**

2 Related work **2**

3 MEC Architecture **3**

4 MEC RAN Abstraction Interface **5**
4.1 Radio Information Interface 6
4.2 Control-plane Interface 7
4.3 Data-plane Interface 8

5 MEC Application Development Framework **8**
5.1 MEC Services 9
5.2 Support Services 10
5.3 Common Services 11

6 MEC Applications **11**
6.1 Active Device Location Tracking 12
6.2 Augmented Reality (AR) Content Delivery 13
6.3 Video Analytics 14
6.4 RAN-aware Content Optimization 16
6.5 Distributed Content and DNS caching 17
6.6 Application-aware Performance Optimization 18

7 Conclusion **20**

List of Figures

1 MEC Architecture 5
2 MEC Interfaces. 6
3 MEC server operation flow 13
4 Active Device Location Tracking 14
5 AR content delivery application 15
6 Video event capture application 16
7 RAN-aware content optimization application 17
8 Distributed content and DNS caching application 18
9 Application-aware Performance Optimization application 19

1 Introduction

Enabling ubiquitous and personalized mobile Internet requires pushing the boundaries of existing network and service infrastructure. Communication services in such situations typically require smart applications encompassing knowledge about the network and users, coupled with content and context information. Especially within the Internet-of-Things (IoT) application areas, (smart) sensing and/or actuating devices and objects are getting involved in human's life generating a tremendous amount of data. This calls for a real-time communication interface to efficiently control and share information across different networks, providers, and geographical areas. In addition, rapid provisioning of new low-latency services requires access to both network control and data plane information with higher flexibility than currently available.

Mobile-Edge Computing (MEC) is emerging as a key technology for the evolution towards 5G to meet the demand to move from a simple bit pipe to a smart service pipe [1–3], and it is a new ETSI standardization initiative [4–8]. While MEC relates, and to some extent relies on software-defined networking (SDN) [9] and ETSI network function virtualization (NFV) [10], which discuss the separation of network functionality from network applications, in order to increase the flexibility of mobile network management, control, and upgrade, it can be seen as an evolution beyond these concepts. In this context, the edge refers to one or multiple Radio Access Networks (RAN) nodes (e.g., eNodeB, Wifi access point) aggregated in a (nano-)data center, and is hierarchically located above the legacy RAN architecture. However, placement of the MEC server(s) and its supported services depends not only on the actual cell deployment (macro-cell, small-cell, heterogeneous) and backhaul network, but also on the service requirements (e.g., low latency) and the distribution of UEs using that service.

MEC allows operators to open their RAN edge service environment to authorized third-parties to rapidly deploy innovative application and service endpoints for the mobile subscribers, enterprises and vertical segments, and network [1]. Such applications can be classified into three categories: device-centric (e.g., online gaming), RAN-centric (e.g., caching), and information-centric (e.g., content-optimization). Following the SDN and NFV principles [11], MEC breaks away with a long-standing assumption of today's cellular systems, by (a) changing the boundary of radio and core network (CN) functions, and (b) providing carriers, users, and applications to control over the network functions themselves and value-added services [12]. This facilitates the development and deployment of complex network services as pieces of software running on the network edge, which carries lots of potentials not only to open the marketplace for network applications, as in the standard SDN case, but to also enable novel applications that require closer proximity to the user.

With MEC, application, services, and content could benefit from a low-latency and high-bandwidth (collaborative) cloud environment in close proximity to the network and user. The distinguishing features of MEC are : (i) proximity to end-users, (ii) direct access to real-time network information, (iii) spatio-temporal context awareness, (iv) mobility support, (v) RAN agnostic, and (vi) network application distribution platform. While MEC follows general SDN principles in terms like abstraction and exposed Application Programming Interfaces (APIs), it goes beyond standard SDN and does not yet have a well defined architecture. One of the key challenges to enable the vision of MEC is the design of an application development framework with its associated services and communication interface. Since the MEC architecture in still under standardization and survey; in this sense, our goal is to propose a suitable MEC architecture that can collaborate with existing radio access network (i.e., LTE/LTE-A) for all MEC use-cases provided by ETSI [1]. In a nutshell, the main contributions of this work are the following:

1. We propose a modular top-down MEC architecture that well integrated with the existing LTE/LTE-A system.

2. We then specify the necessary services of the proposed MEC application framework for the application development environment.

3. Finally, we use the proposed MEC architecture to specify the how all use-cases defined in the ETSI MEC white paper [1] can be applied under proposed architecture.

Last but not least, even we only specify the collaboration between proposed architecture and LTE/LTE-A system; however, our proposed framework can be adapted to different RAN and CN architecture (3G, WiFi) via the modifications on the RAN-dependent part.

The remainder of this paper is organized as follows. Section 2 introduces some related works. Section 3 presents the proposed MEC architecture. MEC communication interfaces and an application development framework are described in Section 4 and 5. Section 6 identifies MEC application interfaces for the five ETSI use-cases. Finally, conclusions and remarks are presented in Section 7.

2 Related work

The MEC is under standardization [4–8], and surveyed [1] by ETSI. Several applications and the advantages of adopting the MEC are proposed recently. Ref. [13] categorizes applications for the deployment at the mobile edge and introduces the benefits for the mobile end users as well as the application service providers. Ref. [14] proposes the Edge Cloud model by augmenting the common cloud data center with service nodes placed at the edges and it shows the advantages in two applications: an accurate and low-latency indoor localization and a scalable-bandwidth video monitoring stream. Ref. [15] provides the REPLISOM architecture which deploys cloud computing resources near IoT service nodes and apply D2D links in order to neutralize the backhauling and routing bottlenecks for IoT networks. Ref. [16] introduces an approach to offload video encoding efforts from mobile devices to external services in existing mobile network architecture and reduce the power consumption of mobile devices. ETSI states the five important use-cases of the mobile edge computing that is targeted for future standardization [1]. Ref. [17] introduces the concept of cache clouds and outline the utility-based mechanism for placing dynamic documents within a cache cloud to improve the performance of the edge cache networks.

Moreover, the MEC is not the only architecture for distributed cloud computing; several similar concepts are proposed to offload the cloud at the edge such as fog computation [18] and cloulet [19]. The fog computing is inspired by the needs of IoT services that require distributed computing and storage and it assumes a completely distributed, multi-layer architecture includes IoT devices, lots of fog computing nodes and the cloud. A detail fog architecture and associated API can be found in [20]. The cloudlet [19] is proposed to have a middle tier within the mobile device - cloudlet - cloud hierarchy in order to bring the cloud closer and provide low-latency, compute-intensive processing. It aims to implement the relevant features as an extension of the open source OpenStack software platform [21].

In our work, we not only propose the suitable MEC application framework but also state how this framework jointly work with the LTE/LTE-A system via the MEC RAN abstraction interface.

3 MEC Architecture

With the emergence of low-latency services, there is a need to study the MEC architecture and identify the modifications to establish communication interfaces with the MEC server. In this paper, we propose a modular MEC server architecture with three main components, namely a MEC RAN abstraction interface, a MEC application development framework, and MEC applications in Fig. 1. The MEC RAN abstraction interface is in charge of establishing communication channels with the underlying network(s) to facilitate efficient control and monitoring of the RAN nodes from the MEC services and applications. It abstracts the details of the network by providing only the necessary information to the MEC application development framework. There are three types of communication channels belong to the proposed MEC RAN abstraction interface:

- **Radio information interface:** provides direct access to real-time radio information through a predefined communication protocol.
- **Control-plane interface:** processes or captures control messages between the RAN and core networks (in LTE this includes X2 application protocol (X2AP) and S1 application protocol (S1AP)).
- **Data-plane interface:** processes data plane messages between RAN and core networks (in LTE this includes X2 and S1 data plane).

The proposed application development framework provides services and high-level APIs for MEC application, and it is composed of four types of service:

- **Common services:** facilitate the usage of the available real-time network and radio information in the local cloud environment. On the control-plane, the Radio Network Information Service (RNIS) provides an abstract view of the network status (e.g., topology, connectivity, load) by extracting the parameters of interest from the RAN with the required level of granularity. On the data plane, the Edge Packet Service (EPS) brings a native IP service end-point to the MEC applications. It acts as a local IP agent performing various network functions, namely IP forwarding, packet encapsulation/decapsulation, and data transcoding.

- **Infrastructure services:** provide physical and/or virtual resources (e.g., computation, storage, network and I/O) with an associated abstraction offered by the service orchestrator (e.g., OpenStack, and OpenStack Heat Orchestrator). Additional features and flexibility may be obtained through a platform service allowing service execution on top of the cloud infrastructure, in an isolated and tenant-based environment.

- **Support services:** provide specific functionalities common to most MEC services. These can be regarded as basic platform services that other, more sophisticated services can utilize in order to facilitate their development. The minimal set of support services include communication service, service discovery and registry, policy and charging service, monitoring service, authentication authorization accounting service, and service-level agreement service. In the following subsection, we only focus on the first two services.

- **MEC services:** provide high-level APIs common to all of the MEC applications and use cases. They allow to build (distributed) network applications based on the abstracted network information on top of the local cloud environment. The considered MEC services are positioning, KPI evaluation and traffic profiling, IP and named data services, event capture, analytic, network status and configuration.

MEC applications are customized control and monitoring programs developed based on the MEC application framework. They can be chained together in the local MEC cloud environment following the NFV service function chaining (SFC) principle to address a particular use-case. They may learn from the past experience and adapt based on cognitive methods to generate knowledge. In particular, these MEC applications may predict future network and user behaviors and forecasting potential solutions according to the history of the collected data with the lowest level of uncertainty. This increases the intelligence of the network and will help automating network operation.

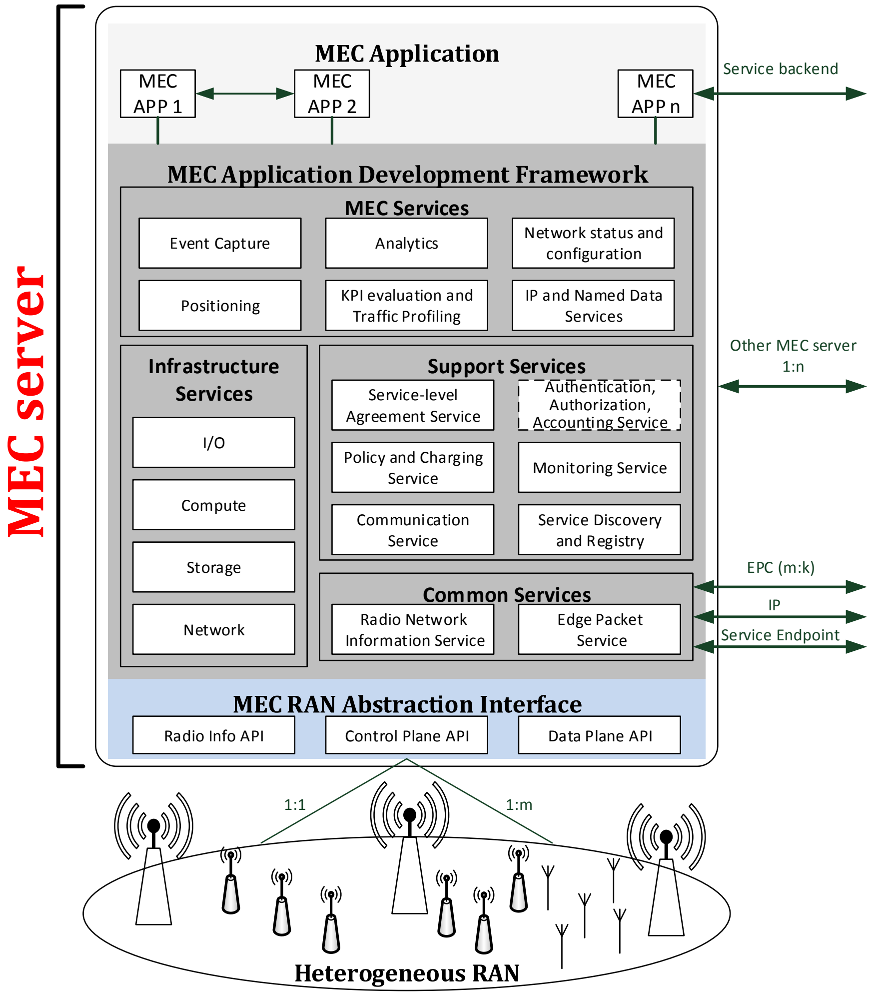

Figure 1: MEC Architecture

4 MEC RAN Abstraction Interface

The RAN abstraction shown in Fig. 1 provides a RAN-specific interface between the MEC and connected physical/virtual network functions. This enables the MEC server to act as a network entity able to communicate with other entities through the RAN-specific control-plane and data-plane interfaces. For example, in sense of LTE/LTE-A network, the MEC server communicates with the Mobility Management Entity (MME) through S1 interface in control plane (S1-C), the serving gateway (S-GW) and Packet data network gateway (P-GW) through the S1 interface in data plane (S1-U) and the eNBs through both X2 control plane (X2-C), S1 and X2 data plane (X2-U and S1-U) in Fig. 2. To enable low-latency service, the

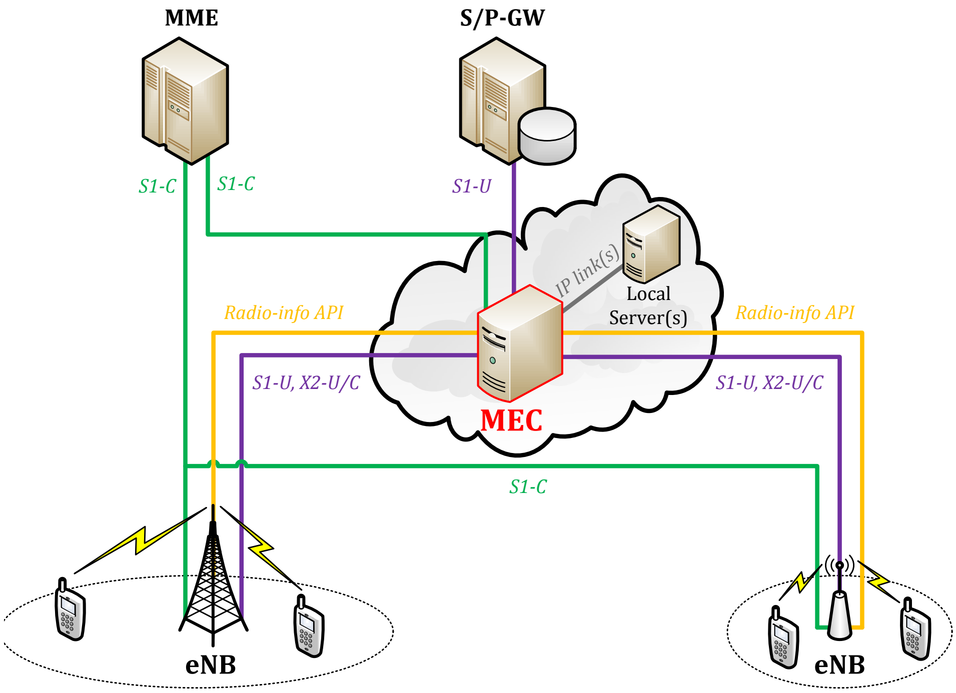

Figure 2: MEC Interfaces.

data plane connection through S1-U/X2-U between the eNB and the S/P-GW must always go through the MEC server. For instance, the user may demand a video that could be sent to the user locally via the MEC server rather via the back-end server. On the other hand, S1-C connection that carries control plane information between MME and eNB is not interrupted by MEC except for the cases of low latency control signaling, e.g., fast handover. In other cases, the legacy S1-C interface between MME and eNB is applied to carry control plane information and these information can be accessed by the MEC server via the S1-C interface between the MME and itself.

The MEC RAN interface handles configuration and status (including context) information in per user, per radio bearer and per carrier basis (i.e. in three dimensions). Configuration information is static or semi-static that can be read or updated, whereas status information is changing overtime but could only be read. The protocol is based on common request/response messages for both configuration and status information that are briefly described in Table 1.

The following subsections describe the above three interfaces in detail.

4.1 Radio Information Interface

This interface is used for collecting information related to UE/eNB lower layer (Layer 1 and Layer 2 in LTE) parameters, IDs and capabilities. The collected information

Table 1: Message on MEC RAN abstraction interface

Message	Field	Usage
Configuration request	Configuration type	Defines the type of configuration, either set or get
	Cell configuration flag	Represents a bit map for the requested cell configuration
	Cell configuration list	A list of cells, identified by their IDs, to request configuration
	UE configuration flag	Represents a bit map for the requested UE configuration
	UE configuration list	A list of UEs, identified by their IDs, to request configuration.
Configuration reply	Cell configuration	A list of cells with the requested configuration report
	UE status	A list of UEs with the requested configuration report
Status request	Status type	Defines the type of status, and can be periodical, one shot, event-driven
	Status period	The status reporting period in terms of Transmission Time Interval (TTI)
	Cell status flag	A bit map for the requested cell status information
	Cell list	A list of cells, identified by their IDs, to request the status
	UE status flag	Represents a bit map for the requested UE status information
	UE status list	A list of UEs, identified by their IDs, to request the status
Status reply	Cell status	A list of cell including the statistic reports
	UE status	A list of UE including the statistic reports

is either of type configuration or status and is stored for further analysis purposes (e.g., network statistics/measurements). Network information obtained through this interface are:

- **UE configuration information:** Public land mobile network (PLMN) ID, Radio network ID (Cell Radio Network Temporary Identifier, C-RNTI in LTE), Downlink (DL)/ Uplink (UL) bandwidth, DL/UL carrier frequency, UE category.
- **UE status information:** Global Navigation Satellite System (GNSS) info (timing/phase), Angle of Arrival (AoA), System Frame Number (SFN), Buffer Status Report (BSR), Timing Advanced, Serving cell Reference Signal Received Power (RSRP)/Reference Signal Received Quality (RSRQ)/ Received Signal Strength Indicator (RSSI), Non-serving cell DL RSRP/RSRQ/RSSI, Device-to-Device (D2D) transmission/reception frequency, D2D Destination info List, UL transmission power.
- **eNB configuration information:** DL/UL radio bearers configuration, Aggregated DL/UL bandwidth, Tracking Area Code (TAC), DL/UL carrier frequency, PLMN ID.
- **eNB status information:** GNSS information, AoA, SFN, DL/UL scheduling information, Wideband and subband channel quality, Buffer occupancy, Timing advanced, Number of active UEs, UL interference level.

4.2 Control-plane Interface

This interface is used to retrieve information related to UE/eNB control-plane, i.e., Upper layer control information (Layer 3 in LTE), S1-C/X2-C parameters and messages, used for network control and monitoring purposes. The information is of type status and it includes:

- **UE status information:** UE capability, Mobility state, Mobility history report, Proximity carrier information, Radio link failure report, Logged measurement report.

- **eNB status information:** eNB X2AP identity, X2AP cause, S1AP cause, Handover restriction list, Physical resource block(PRB) usage per traffic class.

4.3 Data-plane Interface

The data-plane interface is used to capture, analyze, and process data packets and it will be used to provide low-latency data services. This interface communicates through X2-U or S1-U with eNB and S/P-GW in LTE/LTE-A system to retrieve information. The information is further categorized into configuration or status:

- **UE configuration information:** Bearer ID, Bearer type (default or dedicated), Bearer context in terms of QoS (QoS Class Identifier, QCI) and priority (Allocation and retention priority, ARP), Guaranteed bit rate (GBR) and Maximum bit rate (MBR) of GBR bearer, the aggregated maximum bit rate of the UE (UE-AMBR) and the Access Point Name (APN-AMBR) of non-GBR bearer, Packet Data Network (PDN) identity, IP address.

- **UE status information:** Channel quality indication (CQI), Channel state information (CSI), block lock error rate (BLER) of each Hybrid automatic repeat request (HARQ) process, Packet delay, Packet discard rate, Packet loss rate, IP throughput, UE PRB usage.

- **Network configuration information:** GTP tunnel endpoint identifier (S1/S5 interface), dedicated radio bearer (DRB) identity, P-GW address, PDN IP address.

- **Network status information:** Aggregated PRB usage, Delay jitter of specific QCI class, end-to-end delay.

5 MEC Application Development Framework

As mentioned earlier, the MEC application development framework is composed of support services, infrastructure services, common services and MEC services, as in Fig. 1. This framework acts as middleware between the applications and the real radio access network element and signaling and it makes the application developers easily to focus on their specific purpose rather that the details functionality of each radio network element. Inside this framework, all four types of services work jointly to provide the top-down network abstraction and bottom-up value-added information provisioning. The following paragraphs provide a brief description of the MEC services, support services and common services; the infrastructure services are discussed in Section 3.

5.1 MEC Services

These services are used to provide value-added information for both the control-plane and the data-plane, by taking into account UE, RAN, and possible CN elements. They jointly work with the RNIS and EPS in order to have a basic knowledge of the network element information and user traffic characteristics.

To enable the MEC services correctly, the associated parameters (e.g. Update frequency, Operate period, Network element, Target user) of the MEC service are necessary to be indicated explicitly. These parameters are mapped from the arguments of MEC application directly (e.g. Update frequency, Operate period) or indirectly (e.g., Network element, Target user). The arguments of the MEC application are the values to indicate the intended behaviors of the MEC application with high level of abstraction that are introduced in Sec. 6 in more detail. Further, MEC service reports the computed value-added information and responses for information provisioning. We introduce below all six MEC services and the associated parameters:

1. **Positioning**: Computes user geo-location, with different granularity levels, based on the available control-plane measurement information. The location technology (LTE, GNSS) and location method (Distance-based, UL/DL Timing-based, Radio frequency(RF) pattern-based, Satellite-based) are selected based on the specified granularity level.

    - Parameter: Location technology (LTE, GNSS), Location method (Distance-based, Angle-based, Pattern-based, GNSS), Update frequency, Operate period, Network element (explicit element identifier), Target user (specific user identity).

2. **Analytics**: Aims to analyze the control-plane state information from RNIS for value-added information, such as interference map building and network load balancing.

    - Parameter: Target info element (UL/DL air-interference level, Cell capacity, Network load), Update frequency, Operate period, Network element, Target user.

3. **Network status and configuration**: This service enables the MEC server as the RAN controller to adjust RAN control-plane configuration and data plane policy through the data-plane or the control-plane API. The updated configuration is provided by the MEC applications.

    - Parameter: Target bearer (bearer identity), Applied policy/con-figuration (QCI, ARP), Network element, Target user.

4. **KPI Evaluation and Traffic profiling**: Computes specific data-plane KPIs, such as E-UTRAN Radio Access Bearer (E-RAB) accessibility, E-RAB mobility, delay jitter. This service also provides traffic profiles to MEC applications for further usage.

    - Parameter: Target bearer, Target KPI (E-RAB accessibility, E-RAB mobility, Delay jitter), Update frequency, Operate period, Network element, Target user.

5. **Event Capture**: This service analyzes the occurrence of specific events in the data stream (e.g., video stream), and the MEC server acts as the end-node for this stream. The event occurrence flag and some extra reported content are further transferred to the MEC application after successful capture.

    - Parameter: Target bearer, Target event, Extra reported content (Event duration, Reliability), Operate period, Network element, Target user.

6. **IP and Named Data Services**: This service is applied to transport data-plane traffic between MEC and target network element. The legacy IP protocol or Named data networking (NDN) can be applied for packet delivery.

    - Parameter: Target bearer, Protocol configuration (type and header), Operate period, Network element, Target user

5.2 Support Services

Support services are responsible for RAN-independent services, in order to enable the both common service and MEC service functionalities. They comprise communication services, service discovery and registry, policy and charging service, monitoring service, authentication authorization accounting service and service-level agreement service. The communication service, as well as the service discovery and registry, are described in the following paragraphs, as examples.

Communication Services Communication services provide well-defined APIs in order to facilitate communication between the MEC applications and the MEC application development framework as well as internal interaction among the MEC services. The communication can be established via many different proposed architectures depending on vendor design options, e.g., representational state transfer (REST), simple object access protocol (SOAP), publish/subscribe (PUB/SUB), and distributed shared memory (DSM). These architectures provide APIs running on top of hypertext transfer protocol (HTTP), and being protected by security mechanisms against malicious insiders. The type of messages exchanged through these APIs can be common HTTP verbs, i.e., GET, POST, PUT, DELETE, including arguments related to the intended information. This mechanism aids to support abstraction of the underlay network easing communication between the MEC applications and the MEC application framework.

Service Discovery and Registry The service registry identifies the available services, supported by the MEC server to the MEC applications. Common protocol messages provided by the communication services can be used in order to inform the MEC application about the availability of different MEC services, as well as the end-points of the provided ones. The entire component aids the MEC applications to verify if the desired information is available from the underlay network. This mechanism can be implemented as a data base (DB) that includes the holistic information of the available MEC services orchestrating the explicit applications demand.

5.3 Common Services

The common services are in charge of RAN-dependent services. They can be categorized into two main types: radio network information services and edge packet Service that provide specific RAN information for the MEC services functionalities.

Radio network information services (RNIS) RNIS interact through the radio information API and control-plane API with the underlay network in order to provide the requested RAN information. Both APIs are used to collect UE/eNB related content that is stored in a local DB including configuration and status information. Moreover, there are many implementation options regarding the RNIS signaling message mapping to the 3GPP specific ones, supported by the RAN[1], but this is beyond the scope of this paper.

Edge Packet Service (EPS) The EPS provides data-plane traffic transportation between the MEC application and the corresponding data-plane API via traffic routing or packet caching. For the routed traffic, the EPS routes the traffic flow passing through the MEC application framework to reach target users and vice versa. For the cached traffic, the MEC server acts as the end-node network element that communicates with the target users. Further, EPS also interacts through data-plane API in order to have knowledge of the data-plane configuration and information status of each network element.

6 MEC Applications

The MEC applications run on the top of the proposed MEC server architecture in Fig. 1. These applications may belong to a specific use-case depending on the intended functionality. In addition, they can be chained together according to the service bundle definition and can communicate with each other as well as with the external world (e.g. a local or remote server in Fig. 2).

[1]For instance, an agent can act as an intermediate between the RNIS and the RAN in order to make the appropriate signaling message translation according to the 3GPP standards.

Specific arguments are associated with each MEC applications in order to acquire information from the underlay RAN. These arguments include the type of information that is requested, the update frequency (i.e., how often this information will be sent), the operate period (i.e., the measurement/observation period of the requested information), concerned geo-region (i.e., explicit network element) and the target user (i.e., specific target UE).

A specific operation flow is in Fig. 3. First of all, the MEC application and associated arguments are provided through higher-layer API and the communication is enabled by the support services. Then, some indirect mapping is done by RNIS/EPS in order to formulate all required parameters for the MEC services. Afterwards, the MEC services enable the underlay RAN operation (measurement, state and configuration update) through EPS/RNSI and the associated API. The measurement results are reported from RAN through RNIS/EPS to MEC for value-added information computation (i.e. congestion flag, delay jitter). Finally, all raw/value-added RAN information are reported through support service back to the MEC applications.

In following, we provide how the six specific MEC use-cases identified by ETSI to be applied to the MEC applications as well as the related arguments.

6.1 Active Device Location Tracking

This case aims to provide user coordinates with a specific accuracy level which is computed based on the real-time network measurements from active users or eNBs, e.g., GNSS measurement, RSSI and Timing advanced measurement, Round-trip-time (RTT) measurement. The locationing technology and method can be decided based on the accuracy level and the target UE/eNB capability. The MEC application can request the raw measurement information to be provided rather than the user coordinates, since the latter can be computed locally using its locationing algorithm. Further, some extra tracking information could be provided to the MEC application for extra knowledge generation. The Active Device Location Tracking application arguments are as follows:

Arguments: Accuracy level (High/Medium/Low), Report type (User coordinates, Measure information), Extra tracking information (Operator PLMN id, Signal strength, etc.), Update frequency, Operate period, Concerned geo-region, Target user (Category, Access class, id)

The mapping of the MEC application argument to the MEC service parameter is depicted in Fig. 4. The support services interpret the arguments and transport their values to RNIS (Accuracy level, Report type, Tracking information, Geo-region, Target user) and the Locationing service (Update frequency, Operate period) inputs. The RNIS look up in its internal DB for determining the locationing technology and locationing method for target user served by geo-region network element, and these parameters as well as the measurement result will be used by Locationing service for the user coordinates computation. The coordinates are re-

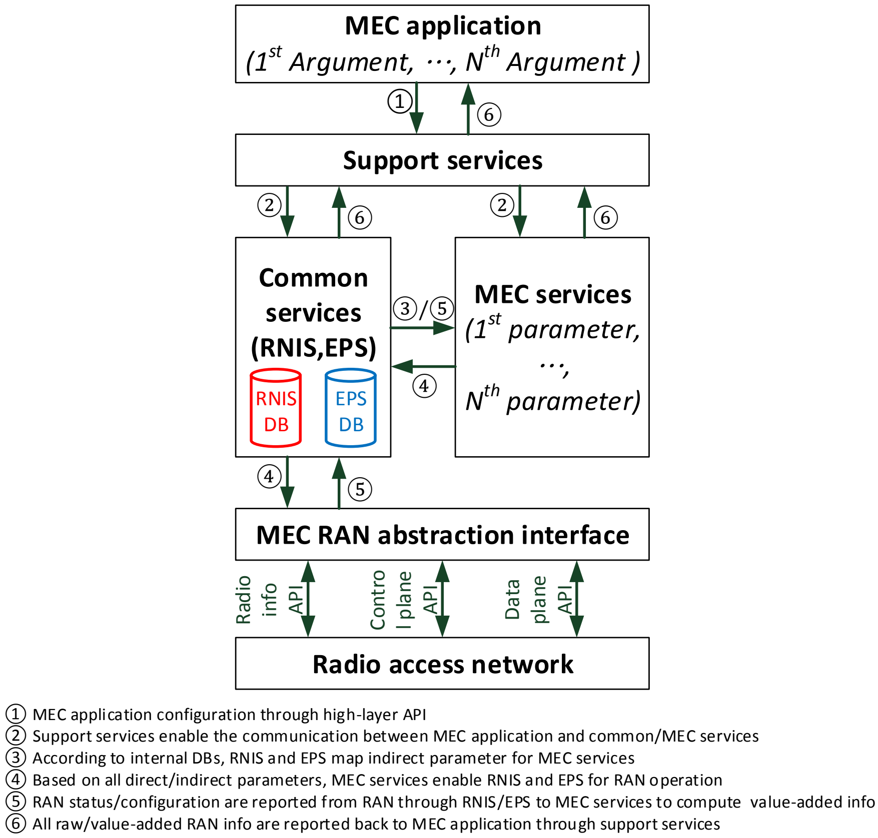

Figure 3: MEC server operation flow

ported back to the MEC application with extra tracking information (PLMN ID, Signal strength) that can be found in the RNIS DB.

6.2 Augmented Reality (AR) Content Delivery

It targets to deliver overlaid AR content based on user geo-location and/or behaviors. The MEC server is acted as a pass-through node for AR content and provides user location as input to the MEC application. The minimum quality of AR content delivery is guaranteed via the AR quality argument that is composed of the guaranteed bit rate (GBR), maximum bit rate (MBR), packet loss rate and priority policy. Moreover, the accuracy of the user locationing is also requested by the MEC application. The arguments of the AR Content Delivery application are defined as:

Arguments: AR quality (GBR, MBR, Loss Rate, Priority), Protocol configuration, Location accuracy, Update frequency, Operate period, Concerned geo-region, Target user

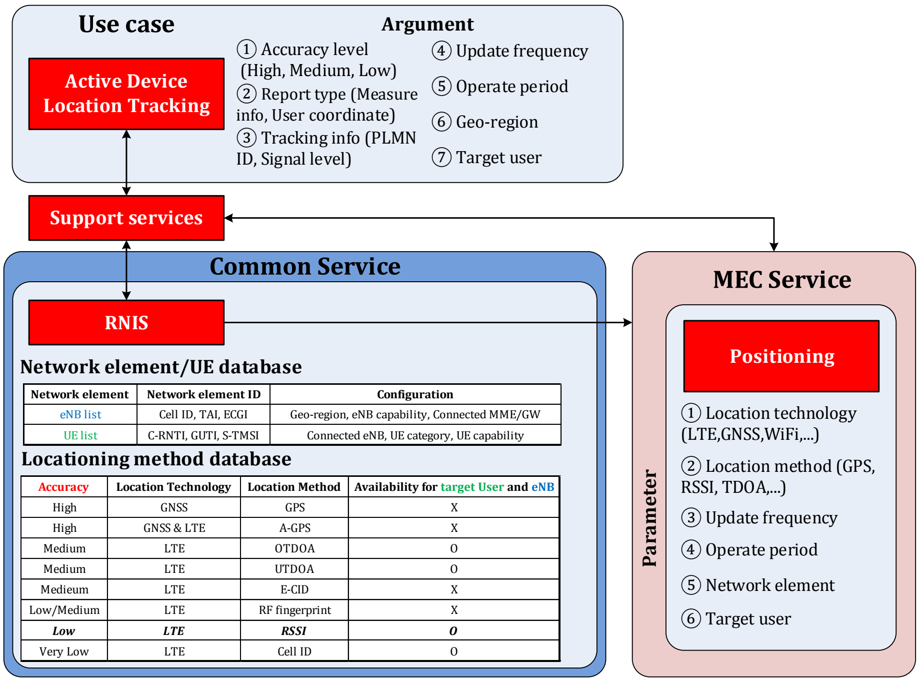

Figure 4: Active Device Location Tracking

Fig. 5 provides the involved MEC services and common services for this MEC application. The RNIS (Location accuracy, Geo-region, Target user) and EPS (AR quality) take the argument values translated by the support services. Afterward, RNIS and EPS will look up in their internal DBs and output the parameters to the three involved MEC services: Locationing, IP and Named Data, and Network status and configuration service.

After receiving the parameters, the Network status and configuration services apply the required QoS parameters to establish a bearer for the AR content. Moreover, the IP and Named Data Services add AR content header with specified protocol configuration for the routed AR content traffic by EPS. Finally, the Locationing services report the user coordinates to the MEC application.

6.3 Video Analytics

This case provides event-based video analysis result to reach less latency and low backhaul bandwidth consumption. In this sense, the original UL video stream is analyzed by the MEC service internally, and only target events with extra report information (event duration, reliability) are reported back to the MEC applications. The MEC application arguments are described as follows, and some arguments are used to guarantee the UL video quality for minimum analytics reliability.

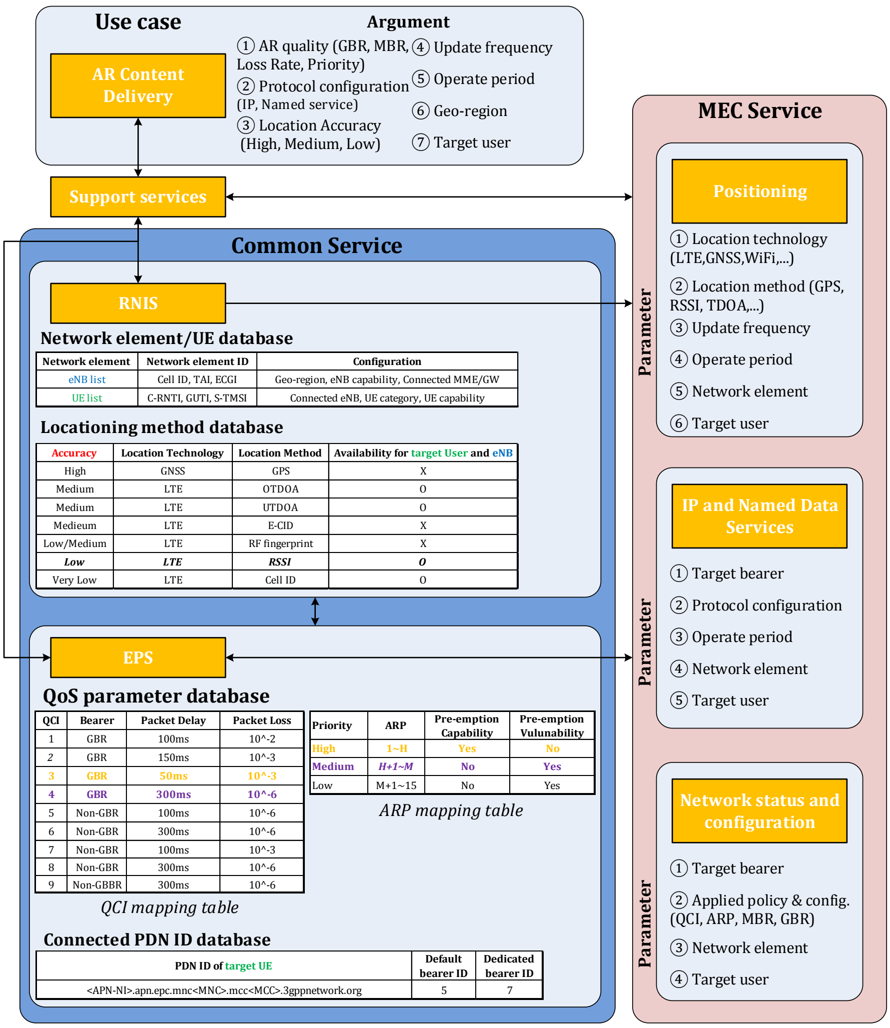

Figure 5: AR content delivery application

Arguments: Target event, Video quality (GBR, MBR, Loss rate, Priority), Report content (Event duration, Reliability), Update frequency, Operate period, Concerned geo-region, Target user

Fig. 6 states the interaction between the involved MEC services and common services. Both RNIS (Geo-region, Target user) and EPS (Video quality) take the translated output (i.e., the application's arguments translated to RNIS/EPS inputs) by the support services and provide the parameters for the correlated MEC services: Event Capture service and Network status and configuration service. The network status and configuration service is applied to establish a video stream bearer for the end-to-end communication between the MEC server and the users. Afterwards, the Event Capture service detects the occurrence of the uplink video flow. The occur-

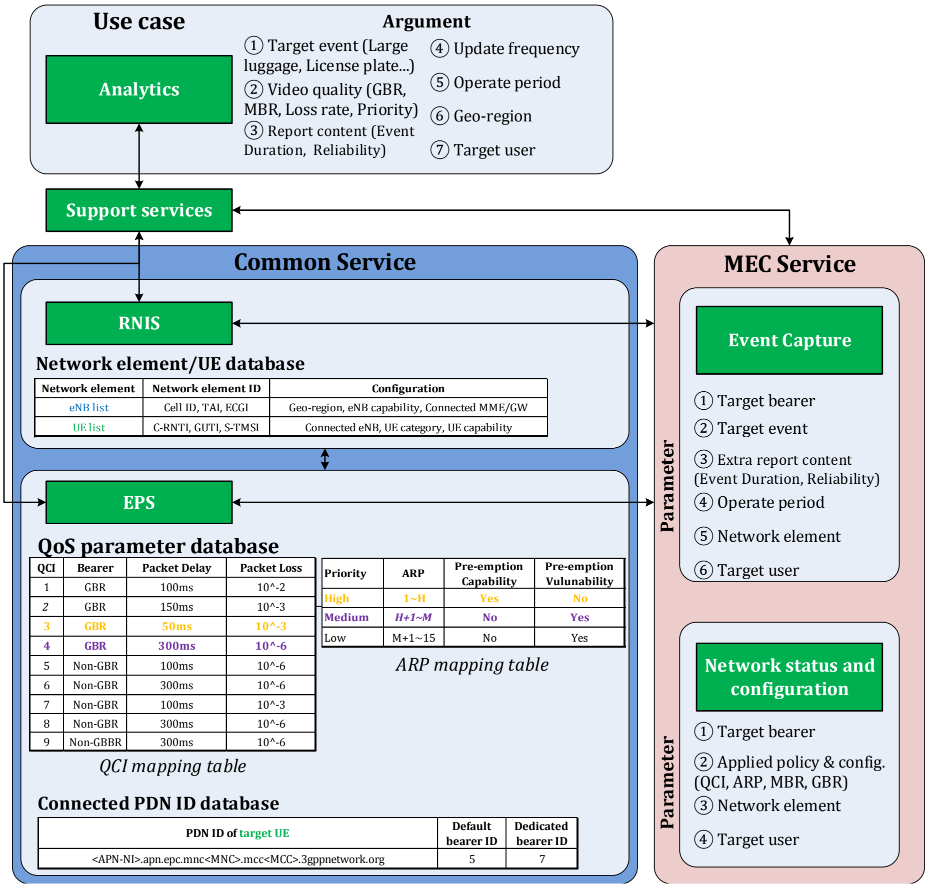

Figure 6: Video event capture application

rence flag and some extra information will be reported back to the MEC application after successful detection.

6.4 RAN-aware Content Optimization

In this case, the MEC application requests the cell and user radio interface status information to deliver a RAN-aware content in order to improve the network efficiency and the quality of experience (QoE). For instance, mobile operators could apply the load status and congestion flag provided by MEC server to reduce the latency and signaling overhead cost. In this sense, this MEC application needs to provide the interesting content identity and the Requested RAN information to the MEC services. The associated arguments are as follows:

Arguments: Content identity(APN ID, MEC, MMC), Requested RAN information (Network load, Packet delay, Interference level, etc.), Update frequency, Operate period, Concerned geo-region, Target user

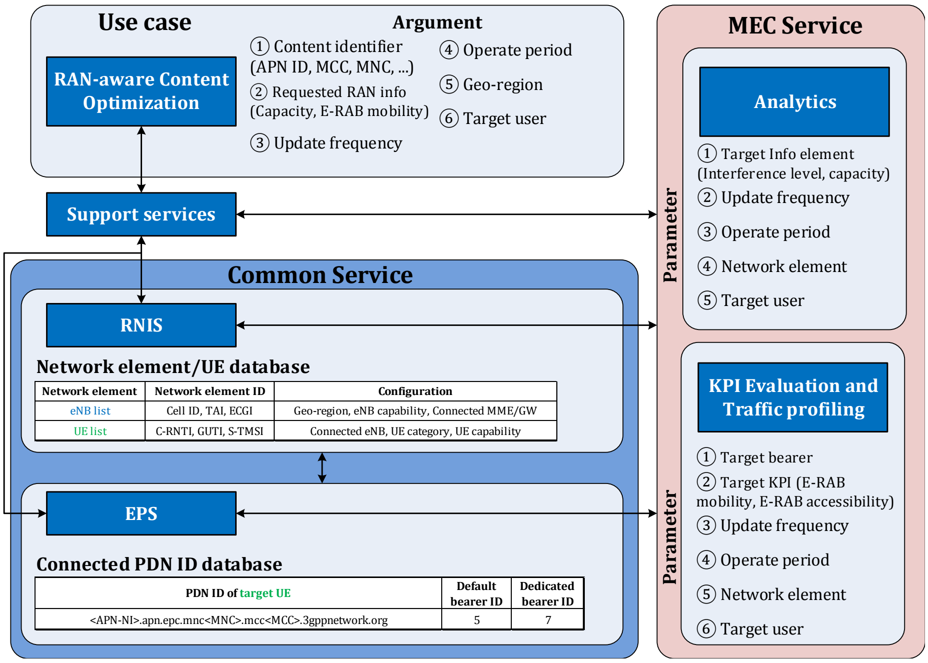

Figure 7: RAN-aware content optimization application

In Fig. 7, the interaction between the common service and involved MEC services is depicted. The input to RNIS (Geo-region, Target user) and EPS (Content identifier) are from the support services. After internal DB look-up at RNIS and EPS, two MEC services will be involve: Analytics and KPI Evaluation and Traffic profiling one. Based on the requested RAN information, the Analytics and the KPI Evaluation and Traffic profiling service are responsible to compute the control-plane and data-plane value-added information and send it back to the MEC application.

6.5 Distributed Content and DNS caching

This case focuses on the distributed content caching for backhaul saving and increasing QoE for content on demand, live streaming, and file/service distribution scenarios. It enables the content to be cached at MEC server for a specified period of time. The MEC application provides the target caching flow identity and content quality argument in order to modify the original non-cached bearer to a new cached bearer. Some associated arguments are as follows:

Arguments: Target caching flow, Update frequency, content quality, Operate period, Concerned geo-region, Target user

In Fig. 8, RNIS, EPS and two MEC services are involved: the IP and Named Data services, and the Network status and configuration service. The RNIS (Geo-region, Target user) and EPS (Target caching flow, Content quality) takes the argu-

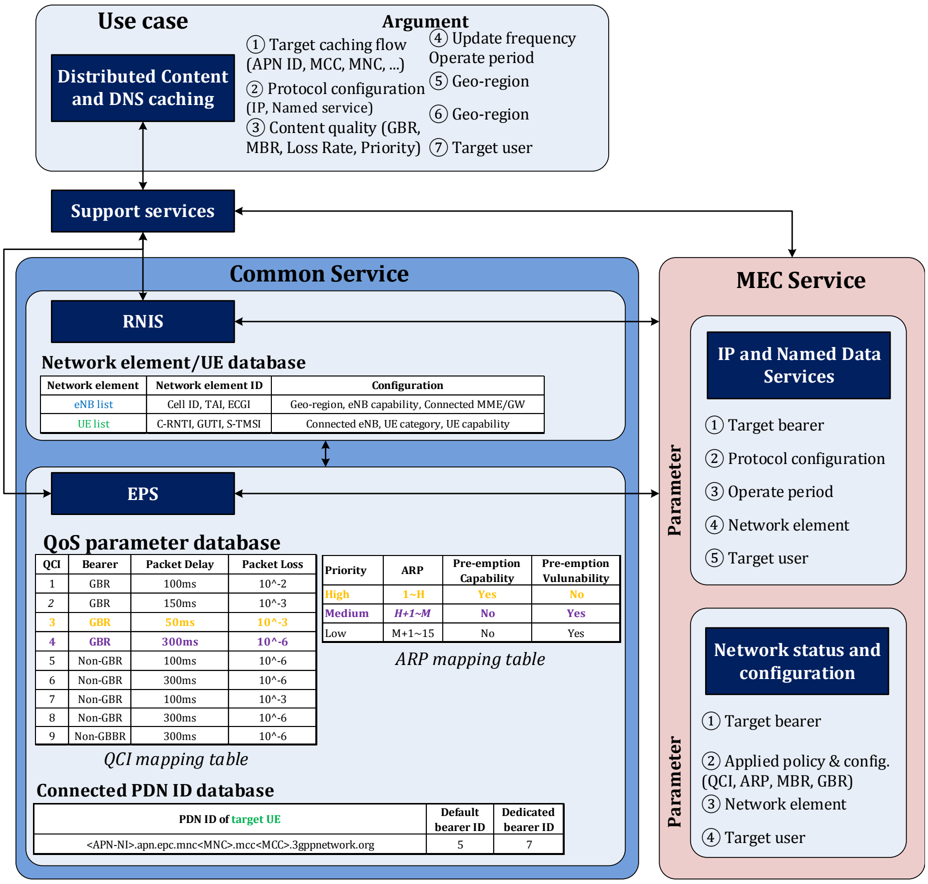

Figure 8: Distributed content and DNS caching application

ment from the support services and look up in their internal DB for the parameters outputted to the two MEC services. The Network status and configuration service modifies the original non-cached bearer into an end-to-end cached bearer from the MEC server to the end users. Afterwards, the IP and Named Data services add the necessary IP or named data header on the cached content for packet delivery performed by EPS.

6.6 Application-aware Performance Optimization

This case aims to improve the network efficiency and QoE, e.g., video stalling time, voice quality, etc., for each higher-layer application (e.g., YouTube, Skype) individually. For instance, operators employ deep packet inspection (DPI) to classify each traffic flow and form the switch flow table rules (e.g., queue prioritization) in order to increase QoE in any IP-based operator network. So, the Application type and Application control protocol arguments are necessary to be identified. Further, two arguments are further identified: Requested RAN information argu-

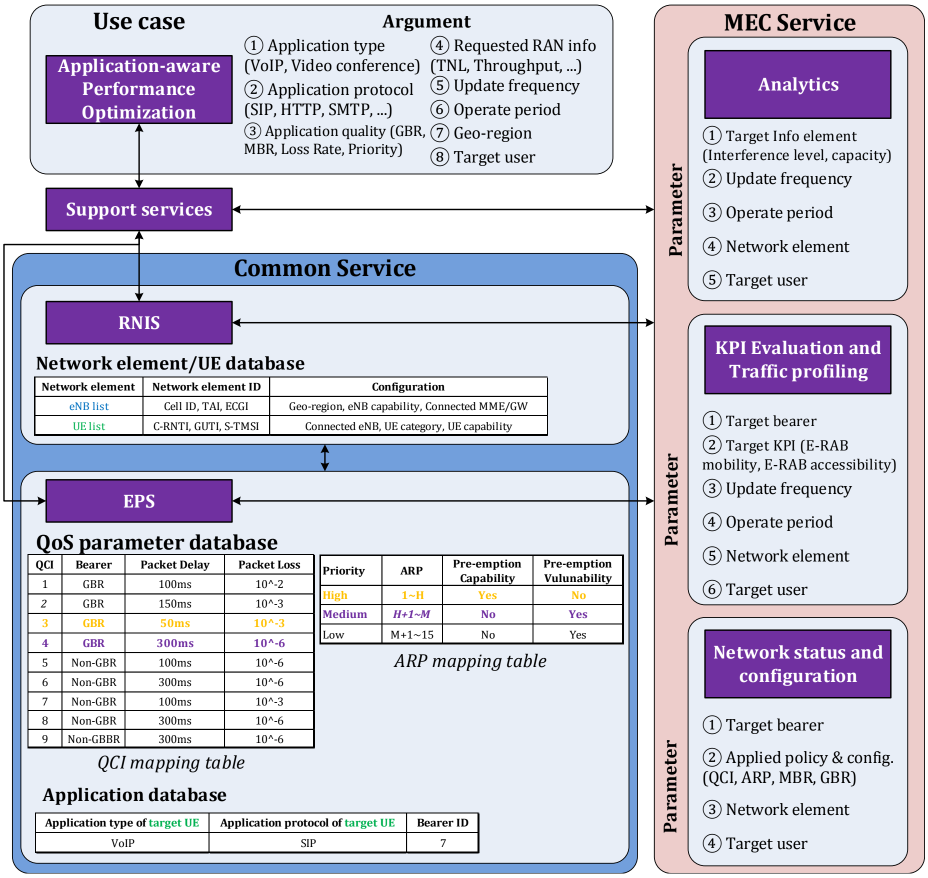

Figure 9: Application-aware Performance Optimization application

ment and Application quality argument. The former argument is used to request RAN information by the MEC application and the latter one is for allocating new QoS policy on the application. The arguments of this case are given as:

Arguments: Application type, Application control protocol, Application quality, Requested RAN information (Throughput, Delay jitter, etc.), Update frequency, Operate period, Concerned geo-region, Target user

In Fig. 9, three MEC services (Analytics, KPI Evaluation and Traffic profiling, Network status and configuration) and two common services (RNIS, EPS) are involved. The RNIS and EPS take the input from the support services and output the parameters to the MEC services after loop-up in their internal DB. Afterwards, the Analytics service and the KPI Evaluation and Traffic profiling service provide the control-plane and the data-plane value-added RAN information that is sent back to the MEC application. Finally, the Network status and configuration service enact the configuration and policy on the target bearer.

7 Conclusion

This paper presents a modular MEC server architecture for LTE/LTE-A that provides a rich network application development environment. It is composed of three main components, (1) MEC RAN abstraction interface in charge of establishing communication channels with the underlying networks, (2) MEC application development framework provides services and high-level APIs for application development, and (3) MEC application offering value-added network service. Then, the proposed architecture is used to analyze the ETSI MEC use-cases, and derive the required network information and the mapping of the MEC services. The analysis reveals the interplay between MEC, RAN and cloud to enable a slicable network for the future 5G system. Going forward, we plan to build a MEC proof-of-concept for RAN-aware video optimization and low-latency service for IoT leveraging OpenAirInterface software implementation of LTE/LTE-A system.

Acknowledgment

Research and development leading to these results has received funding from the European Framework Program under H2020 grant agreement 671639 for the COHERENT project.

References

[1] M. Patel, J. Joubert, J. R. Ramos, N. Sprecher, S. Abeta, and A. Neal, "Mobile-edge computing," *ETSI*, 2014.

[2] Intel and Nokia-Siemens, "Increasing mobile operators' value proposition with edge computing," Intel and Nokia-Siemens, Tech. Rep., 2013.

[3] R. E. Hattachi, J. Erfanian, and B. Daly, *NGMN 5G Initiative White Paper*. NGMN, 2015.

[4] ETSI MEC ISG, "Mec001: Mobile edge computing (mec) terminology," ETSI, Tech. Rep., 2015.

[5] ——, "Mec002: Mobile edge computing (mec) technical requirements," ETSI, Tech. Rep., 2015.

[6] ——, "Mec003: Mobile edge computing (mec) framework and reference architecture," ETSI, Tech. Rep., 2015.

[7] ——, "Mec004: Mobile edge computing (mec) service scenarios," ETSI, Tech. Rep., 2015.

[8] ——, "Mec005: Mobile edge computing (mec) poc framework," ETSI, Tech. Rep., 2015.

[9] D. Kreutz, F. Ramos, P. Esteves Verissimo, C. Esteve Rothenberg, S. Azodolmolky, and S. Uhlig, “Software-defined networking: A comprehensive survey,” *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2015.

[10] ETSI, “Network Functions Virtualisation (NFV), White paper,” Tech. Rep., 2014.

[11] R. Schooler, “Transforming networks with nfv & sdn,” *Intel Architecture Group*, 2013.

[12] N. Nikaein, E. Schiller, R. Favraud, K. Katsalis, D. Stavropoulos, I. Alyafawi, Z. Zhao, T. Braun, and T. Korakis, “Network store: Exploring slicing in future 5g networks,” in *Proceedings of the 10th International Workshop on Mobility in the Evolving Internet Architecture*, ser. MobiArch ’15, 2015.

[13] M. T. Beck, M. Werner, S. Feld, and S. Schimper, “Mobile edge computing: A taxonomy,” in *Proc. of the Sixth International Conference on Advances in Future Internet*. Citeseer, 2014.

[14] H. Chang, A. Hari, S. Mukherjee, and T. Lakshman, “Bringing the cloud to the edge,” in *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on*. IEEE, 2014, pp. 346–351.

[15] S. Abdelwahab, B. Hamdaoui, M. Guizani, and T. Znati, “Replisom: Disciplined tiny memory replication for massive iot devices in lte edge cloud,” 2015.

[16] M. T. Beck, S. Feld, A. Fichtner, C. Linnhoff-Popien, and T. Schimper, “Mevolte: Network functions for energy-efficient video transcoding at the mobile edge,” in *Intelligence in Next Generation Networks (ICIN), 2015 18th International Conference on*. IEEE, 2015, pp. 38–44.

[17] L. Ramaswamy, L. Liu, and A. Iyengar, “Cache clouds: Cooperative caching of dynamic documents in edge networks,” in *Distributed Computing Systems, 2005. ICDCS 2005. Proceedings. 25th IEEE International Conference on*. IEEE, 2005, pp. 229–238.

[18] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, “Fog computing and its role in the internet of things,” in *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*. ACM, 2012, pp. 13–16.

[19] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, “The case for vm-based cloudlets in mobile computing,” *Pervasive Computing, IEEE*, vol. 8, no. 4, pp. 14–23, 2009.

[20] F. Bonomi, R. Milito, P. Natarajan, and J. Zhu, “Fog computing: A platform for internet of things and analytics,” in *Big Data and Internet of Things: A Roadmap for Smart Environments*. Springer, 2014, pp. 169–186.

[21] K. Ha and M. Satyanarayanan, “Openstack++ for cloudlet deployment,” *School of Computer Science Carnegie Mellon University Pittsburgh*, 2015.