ETSI GS MEC-IEG 004 V1.1.1 (2015-11)

Mobile-Edge Computing (MEC); Service Scenarios

Disclaimer

This document has been produced and approved by the Mobile-Edge Computing (MEC) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference

DGS/MEC-IEG004

Keywords

ETSI

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

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

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Contents

Intellectual Property Rights

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

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

Foreword

This Group Specification (GS) has been produced by ETSI Industry Specification Group (ISG) Mobile-Edge Computing (MEC).

Modal verbs terminology

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

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

Introduction

Mobile-Edge Computing provides IT service environment and cloud-computing capabilities within the Radio Access Network (RAN) in close proximity to mobile device. MEC aims to reduce latency, ensure highly efficient network operation, service delivery and ultimate personal experience. The term "Edge" in this context means the radio base station itself (eNodeB, RNC, etc.), and servers within the radio network (e.g. at "aggregation points"). The presence of MEC server at the edge of the RAN allows exposure to real-time radio and network information (such as subscriber location, cell load, etc.) that can be leveraged by applications and services to offer context-related services; these services are capable of differentiating the mobile broadband experience.

The Mobile-Edge Computing environment creates a new value chain and an energized ecosystem, which in turn can create new opportunities for mobile operators, application and content providers whilst enabling them to play complementary and profitable roles. Based on innovation and business value, this value chain will allow all players to benefit from greater cooperation and better monetize the mobile broadband experience.

Mobile Operators can open up the radio network edge to third-party partners, allowing them to rapidly deploy innovative applications and services towards mobile subscribers, enterprises and other vertical segments.

For application developers and content providers, the RAN edge offers a service environment with ultra-low latency and high bandwidth as well as direct access to real-time radio and network information. Mobile-Edge Computing allows content, services and applications to be accelerated, increasing responsiveness from the edge. The customer experience can be proactively maintained through efficient network and service operations, based on insight into the radio and network conditions.

Network vendors and technology providers can provide RAN equipment and base stations enhanced with cloud-computing capabilities, by offering more powerful and flexible network elements able to satisfy the increasing needs of the communication world.

The goal of the present document is to describe some service scenarios that can be delivered through the use of Mobile-Edge Computing. These serve as examples of how MEC can stimulate innovative services and applications that would create a better quality of experience for the end user. MEC will enable a large number of new kinds of applications and services for multiple sectors (such as consumer, enterprise, health, etc.).

1 Scope

The present document introduces a number of service scenarios that would benefit from the introduction of Mobile-Edge Computing (MEC) technology. Mobile-Edge Computing enhances the Mobile Access Network with MEC servers that enable applications to run in the mobile edge. Due to the proximity of the server into the Access network, there is a latency reduction of delivering such applications or services. Such applications can expose information that can be used to optimize the network and services, reduce latency, and support creating personalized and contextualized services. Moreover, new and innovative analytics services enable the operator to monitor usage and service quality. Internet of Things (IOT) applications or enterprise communications can benefit greatly from MEC, as it allows service delivery in close proximity to the actual terminal devices.

The focus of the present document is to introduce or provide a non-exhaustive set of service scenarios. It is not the intent nor does the present document provide any requirements.

2 References

2.1 Normative references

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

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

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

The following referenced documents are necessary for the application of the present document.

Not applicable.

2.2 Informative references

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

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

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

Not applicable.

3 Definitions and abbreviations

3.1 Definitions

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

analytics: processing which makes use of data to provide actions, insights and/or inference

augmented reality: live direct or indirect view of a physical, real-world environment whose elements are enhanced, or supplemented, by computer-generated sensory

Internet of Things (IoT): device, machine or entity connected to any other device, machine or entity through a network

vehicle: human transport including automobile, car, truck, plane, train, motorcycle

3.2 Abbreviations

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

DSRC	Dedicated Short Range Communications
IoT	Internet of Things
IP-PBX	Internet Protocol Private Branch Exchange
LAN	Local Area Network
LTE	Long Term Evolution
MEC	Mobile-Edge Computing
PBX	Private Branch Exchange
QoE	Quality of Experience
RAN	Radio Access Network
RNC	Radio Network Controller
TCP	Transmission Control Protocol
WLAN	Wireless Local Area Network

4 Service scenarios

4.1 Intelligent video acceleration service scenario

4.1.1 Description of scenario

Eliminating mobile content delivery inefficiencies by guiding sources as to real-time network capacity.

A Radio Analytics application located at the RAN provides the video server with an indication on the throughput estimated to be available at the radio downlink interface. The information can be used to assist TCP congestion control decisions and also to ensure that the application-level coding matches the estimated capacity at the radio downlink.

4.1.2 Motivation

Improve the end user's Quality of Experience (QoE) by reducing the content's time-to-start as well as video-stall occurrences, and guarantee maximum utilization of the radio network's resources.

4.1.3 Problem statement

The behaviour of the Transmission Control Protocol (TCP), which assumes that network congestion is the primary cause for packet loss and high delay, can lead to the inefficient use of a cellular network's resources and degrade application performance and user experience.

The root cause for this inefficiency lies in the fact that TCP has difficulty adapting to rapidly varying network conditions. In cellular networks, the bandwidth available for end devices can vary by an order of magnitude within a few seconds due to changes in the underlying radio channel conditions, caused by the movement of devices, as well as changes in system load when other devices enter and leave the network.

4.1.4 Relation to MEC

Figure 1 shows an example of the intelligent video acceleration service scenario. In this scenario, a Radio Analytics application, which resides in a MEC server, provides the video server with a near real-time indication on the throughput estimated to be available at the radio downlink interface.

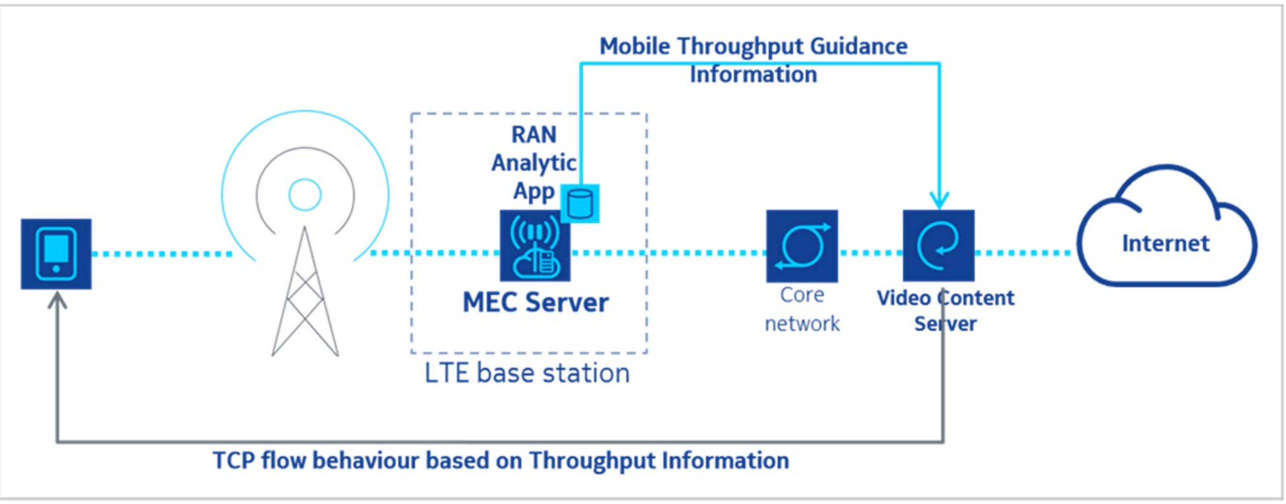

Figure 1: Intelligent Video Acceleration

The video server may use this information to assist TCP congestion control decisions, for example in selecting the initial window size, setting the value of the congestion window during the congestion avoidance phase, and adjusting the size of the congestion window when the conditions on the "radio link" deteriorate. In other words, with this additional information, TCP does not need to overload the network when probing for available resources, nor does it need to rely on heuristics to reduce its sending rate after a congestion episode.

This information may also be used to ensure that the application-level coding matches the estimated capacity at the radio downlink. The aim of all of these improvements is to enhance the end user's quality of experience by reducing the content's time-to-start as well as video-stall occurrences, and to guarantee maximum utilization of the radio network's resources.

Mobile-edge Computing allows the Radio Analytics application to be deployed on top of platforms implemented by different vendors and across multi-operator networks. This ensures efficient utilization of the network resources in addition to enhanced quality of experience for the vast majority of end users.

4.2 Video stream analysis service scenario

4.2.1 Description of scenario

Consider a video based monitoring system such as vehicle license plate recognition for example to monitor vehicles entering and exiting an area of the city, car parks, for security purposes, etc. License plate information can be captured and sent to a cloud based monitoring service. It may be useful to extract other information from such video streams.

4.2.2 Motivation

Moving the video analysis away from the video camera reduces the cost of such cameras, especially when numerous cameras need to be deployed. Furthermore, performing the analysis locally, i.e. close to the mobile base station, mitigates the need to transmit high data video streams when only small pieces of information are required to be extracted from these video streams.

4.2.3 Problem statement

Currently, video based monitoring either requires video streams to be sent to a server or for video processing to be done at the same site as the camera. Both methods are costly and inefficient when compared to processing video data at a MEC server in order to extract meaningful data from video streams. The valuable data can then be transmitted to the application server without the need to transport high data rate video streams.

4.2.4 Relation to MEC

The use of a MEC server is highly advantageous in allowing flexibility on the analysis performed on video streams, especially compared with video processing done at source (at the camera site). Additionally, it saves from transporting potentially numerous high data video streams through the core network to cloud based services.

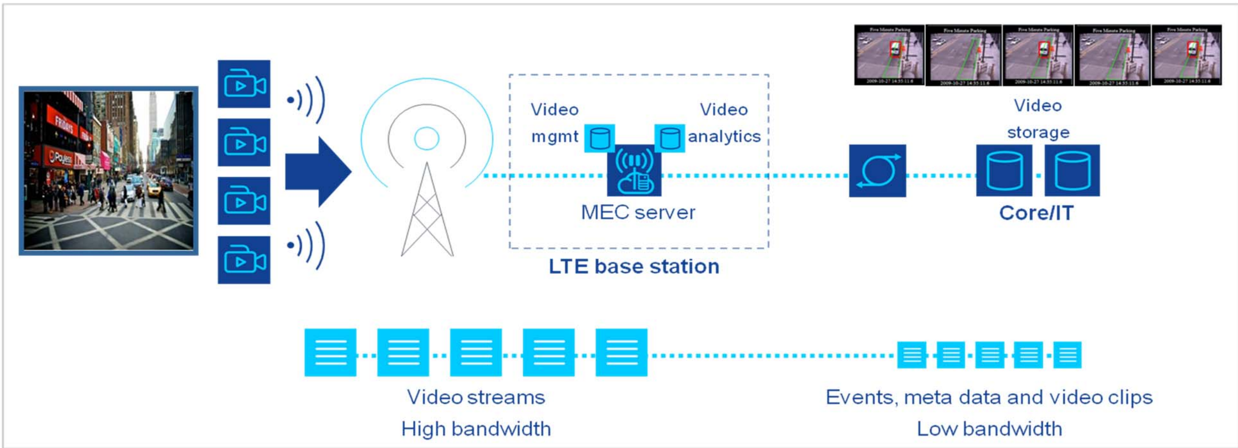

Figure 2: Video stream analysis

4.3 Augmented reality service scenario

4.3.1 Description of scenario

A visitor to a museum, art gallery, city monument, music or sports event holds their mobile device towards a particular point of interest with the application related to their visit activated (i.e. the museum application). The camera captures the point of interest and the application displays additional information related to what the visitor is viewing.

4.3.2 Motivation

Augmented reality can enhance the experience of a visitor to a museum or another point of interest. Augmented information pertaining to a point of interest is highly localized and thus hosting the information locally is advantageous compared with hosting in the cloud.

4.3.3 Problem statement

Augmented reality service requires an application to analyse the output from a device's camera and/or a precise location in order to supplement a user's experience when visiting a point of interest by providing additional information to the user about what they are currently experiencing. The application needs to be aware of a user's position and the direction they are facing, either through positioning techniques or through the camera view, or both. After analysing such information, the application can provide additional information in real-time to the user. If the user moves, the information needs to be updated.

4.3.4 Relation to MEC

The use of a MEC server is highly advantageous since augmented reality information is highly localized. Additionally, the processing of user location or camera view can be performed on a MEC server rather than on a more centralized server. There may be a need to update information at a fast rate, depending on how the user moves, and the context in which the augmented reality service is used (e.g. in an art gallery, exhibits are positioned only a few metres apart and each piece is supplemented with additional text on the artist, the interpretation of the artwork, etc.). In other words, augmented reality data requires low latency and a high rate of data processing in order to provide the correct information to the user's device depending on the location and orientation of the user.

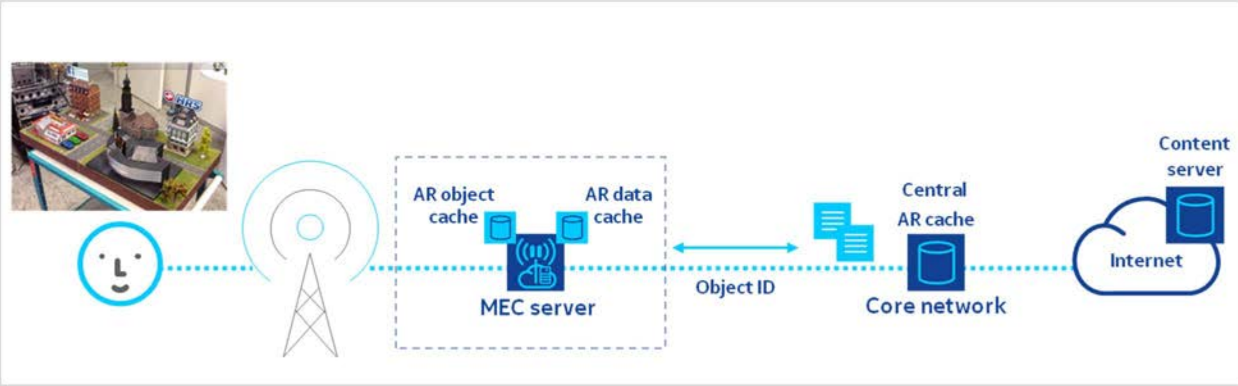

Figure 3: Augmented reality

4.4 Assistance for intensive computation

4.4.1 Description of scenario

In order to maximize battery life and simplify a device or sensor in order to make it low cost, intensive computation and decision making may be offloaded to the network. In some cases, such devices may also require information from other remote sources which is fed into the data processing required to do some decision making which is then fed back to the device. One example could be a robotic device that encounters an object and is required to perform image recognition. The image details are communicated to the server which returns a result. The robot might be instructed to perform an action on the object, e.g. if the object is a ball, kick it with your right leg. Gaming, environmental sensors and security applications could be other examples where offloading computations can be useful to enhance performance.

4.4.2 Motivation

Offload computationally intensive data processing which may require input from multiple sources. This can improve performance of a device with low processing power and also improve battery performance.

4.4.3 Problem statement

A device or sensor is made to be as low cost as possible but is required to remain operational for a long period of time. Such a device may also require further instruction or feedback based upon the information it feeds to the application or service. As a result of the low cost, computational power is often sacrificed. In some cases such computation requires input from other remote sources. Performing such computation off-board the device and within an entity where high performance processing is available can increase the battery life of remote devices and make them low-cost where such criteria is an important (e.g. sensors, meters, security systems where certain actions are taken based upon detected event). This type of computation offload may require applications to be redesigned to enable some features of the application to reside in the device and other aspects of the application to reside on a network resource. Some offloaded computation may require very low processing times with low latency to return back to the remote device the result of computation.

4.4.4 Relation to MEC

A MEC server can be used to host high performance computation capabilities which can take information from multiple sources. Such computation can be done in very short timescales, and the result fed back to remote devices which may require information to perform a further action. Such deployment reduces the need for a remote device to not only perform intensive data processing, but also reduces the need for a remote device to receive information from multiple sources in order to perform some meaningful computation.