MOVING SERVICE FUNCTIONS INTO EDGE CLOUD ENVIRONMENTS ON THE FLY

Kay Hänsge, Roman Szczepanski, Hans Einsiedler, Michael Maruschke*

22. VDE/ITG Fachtagung Mobilkommunikation – 9th/10th May 2017 – Hochschule Osnabrück

T · · * HfTL

LIFE IS FOR SHARING.

INTRODUCTION

- Mobile-Edge Computing (MEC) will be one of the key elements of the next generation communication infrastructure. The 5G infrastructure will not only address a new air interface, but will be a paradigm change of telecommunication.

- Drivers of this change are the vertical industrial areas (e.g., automotive, e-health, smart grids, smart cities, industrial internet) and of course the needs and expectations of end-customers and users.

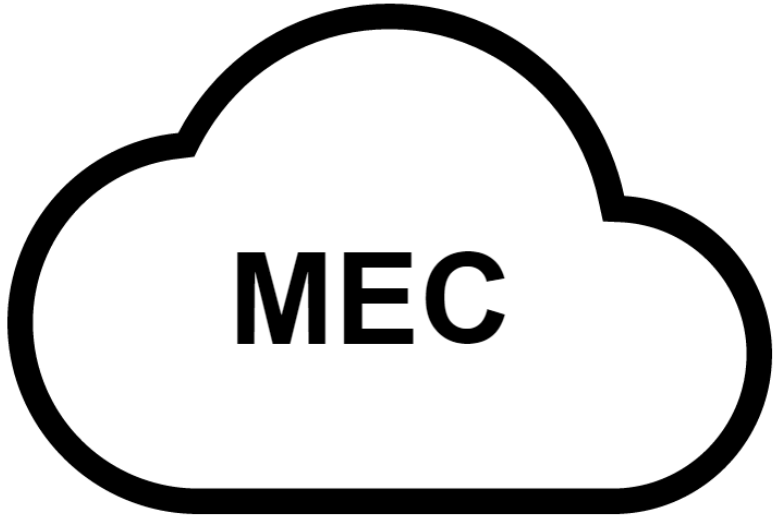
- Side Fact: ETSI changed the name from Mobile-Edge Computing to Multi-Access Edge Computing

INTRODUCTION

- Virtualisation, Software-Defined Networking (SDN), and cloud concepts will enforce this change. The future telecommunication resources will consist of connectivity, storage capacity, and processing power in a dynamic and possible virtualized manner.
- This flexibility enables the implementation of virtualised networks – network slices – to satisfy the requirements of the vertical industries and the users.
- Within a network slice the control plane and the data plane will be optimised according to the mentioned vertical and user requirements – such as deploying network services within context of MEC.

MEC - BASICS

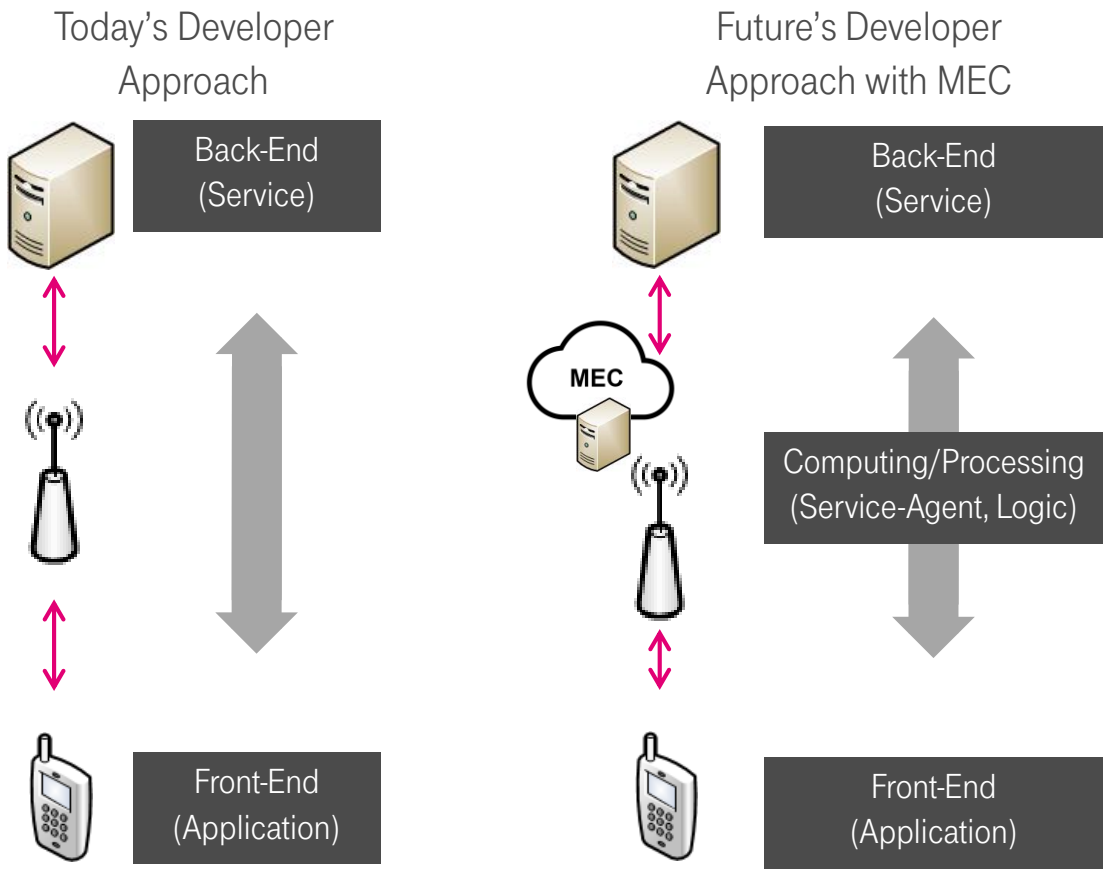
- **MEC is an approach to reduce**
  - wireless and mobile network service latency (not on air Interface),
  - bandwidth usage through the backhauling and core network.
- **It focuses use cases and services which are using**
  - localization-based services,
  - real-time critical communications, and
  - current radio network information.



- **The running services are distributed on the providers' network infrastructure.**
- **New and upcoming business models can be derived by offering such platform/infrastructure features.**

Image: Roselin Christina S. from the Noun Project

MEC - NEW APPROACHES TO DEVELOP APPLICATIONS



Back-End can be an OTT service provider (chosen by the application developers).

MEC Platform is provided by the Operator closer to the Access and consumer.

Front-End Application remains on the mobile terminal.

MEC - SERVICE SCENARIOS

Intelligent video acceleration
- content delivery will be optimized by guiding the data sources

Video stream analysis
- processing of the video will be done at the edge and extracted information then pushed to services

Augmented reality
- where additional information is displayed based on the currently viewed point of interest

Assistance for intensive computation
- where the processing will be done in the network and data is injected by the customers end device

Enterprise deployment of MEC
- where fixed network elements are replaced by mobile equipment

Connected vehicles
- context messages can be distributed in an area for efficiency and safety purposes

IoT gateway
- where enormous amount of data is aggregated which and then forwarded as a single bulk of data

Follow the User
- where a service may follow the consumer while he is moving to offer always best conditions

Private & Local Processing
- covering the privacy and secure processing of data

Local Processing of I4.0 productions
- control applications for industrial machines may run on an on-premises network node

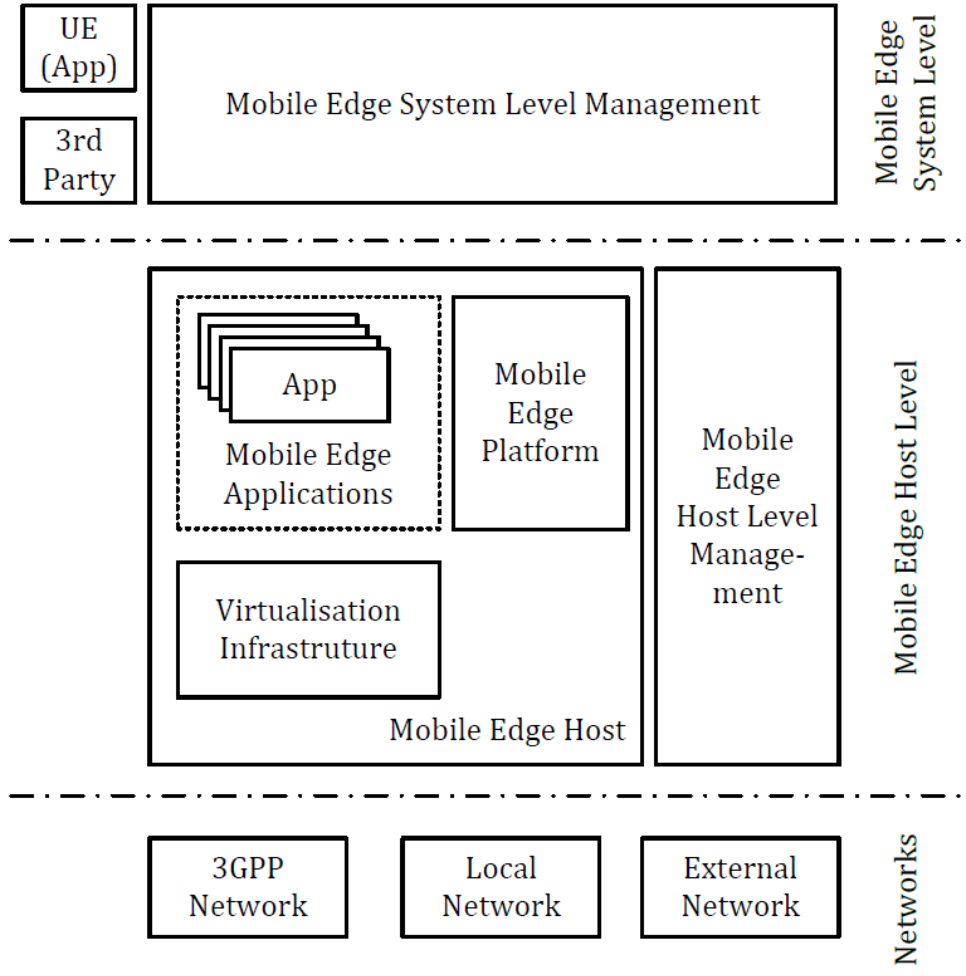
smart grid and energy management
- always fast and secure control

... and many more ...

Starting Source: ETSI, “Mobile-Edge Computing (MEC); ETSI, “MEC Service Scenarios,” ETSI, GS MEC 004

MEC - THE ETSI-BASED FRAMEWORK

- The **Mobile Edge System Level** offers orchestration and includes the Operations Support System (OSS). The applications of the User Equipment (UE) or 3rd Party interact with the Mobile Edge System Level Management to request MEC functionality for a specific service.
- The **Mobile Edge Host Level** comprises the Mobile Edge Host and the associated Management entity. The Mobile Edge Platform and the Virtualisation Infrastructure are located in the Mobile Edge Host, hosting the actual Edge Applications.
- The **Networks** area offer the required connectivity for the data plane communication (such as 3GPP, local, or external networks).



Source: ETSI, "Mobile-Edge Computing (MEC); Framework and Reference Architecture," ETSI, GS MEC 003

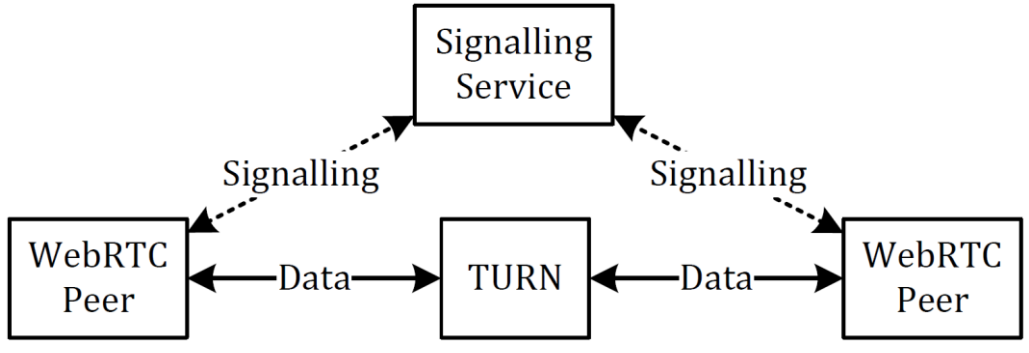
MOVING CONTROL POINTS CLOSE TO THE CONSUMPTION POINT HOW COULD IT BE DONE?

- Executing services as close as possible (at base station or at a nearby connected small data center) to the service consumption point is an upcoming challenge for the future networks.
- Delay critical applications benefit from fast response times and hence, offer a better service for the consumers.
- Placement and Migration procedures need to be performed by an orchestrating entity which monitors the hosting infrastructures.

Pre-Deployment	• Placement, somewhere in the cloud
On-Demand	• Placement at session setup
Live-Migration of an Instance	• Migrate full-fledged VM • Migrate Application-Containers
State Migration	• Extract & Inject Session States

WEBRTC - AN OUTLINE

- WebRTC enriches web browsers with HTML5-based real-time communication functions offering High Definition (HD) audio and video exchange bidirectional, without any further installation of plugins or applications.
- Figure shows the typical and generalised data paths where two Peers are exchanging user data and signalling via a service.
- TURN-Server, which acts as a communication control point as well as levering firewall or gateway issues.

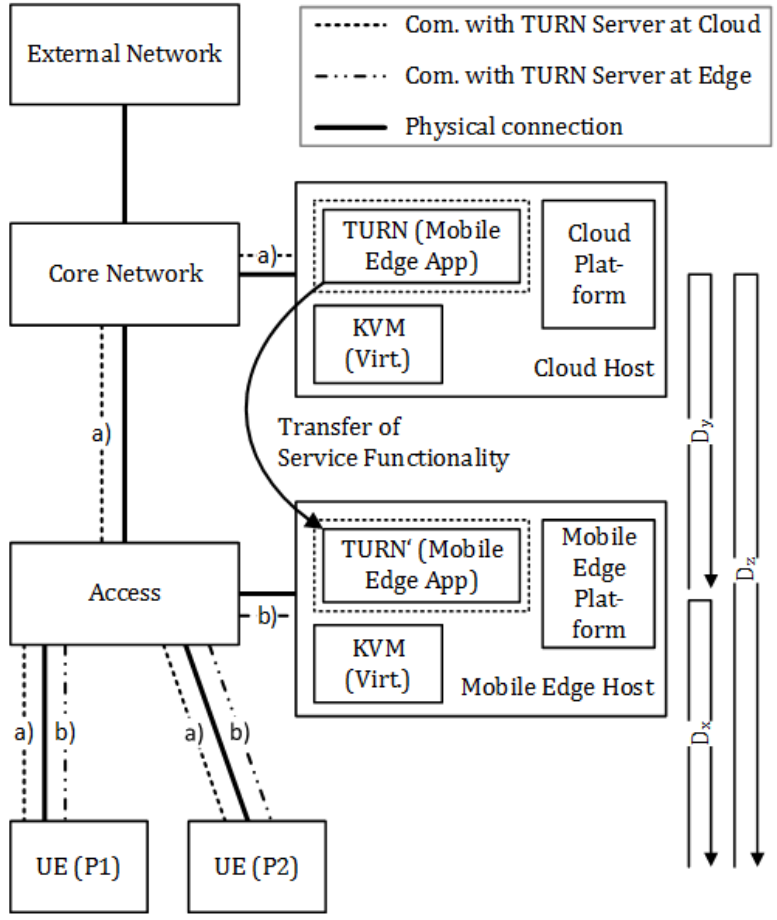


Traversal Using Relays around NAT

Source: http://webrtc.org

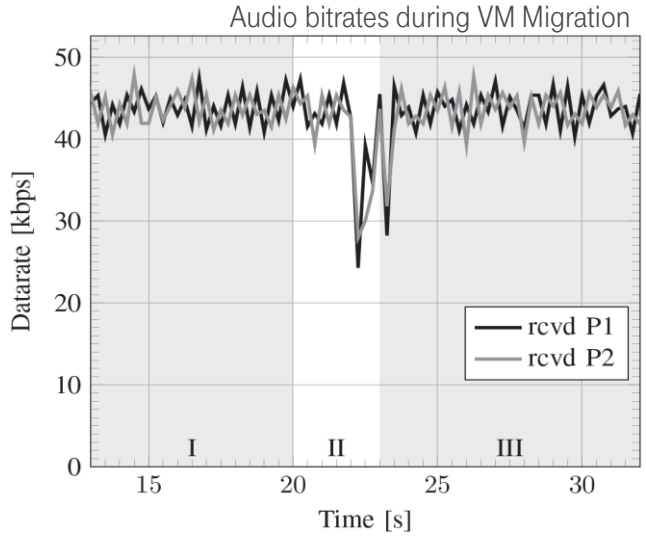
WEBRTC TURN SERVER MOVEMENT SCENARIO

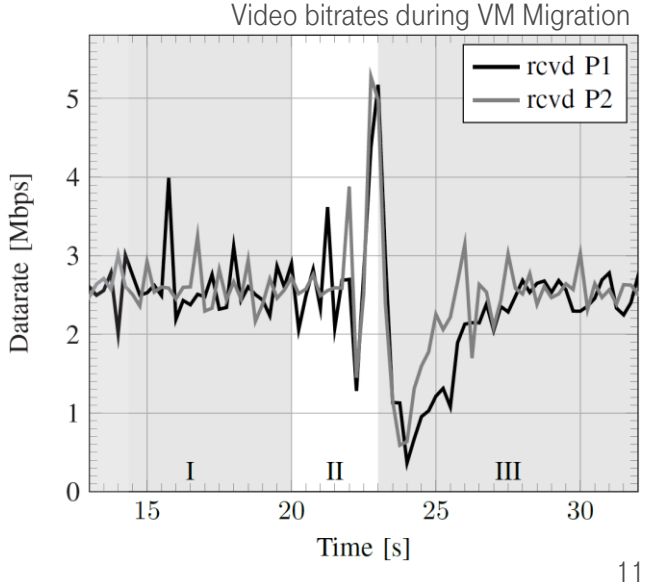
- Two users are currently travelling in the same foreign country and want to communicate. Hence, they are associated and connected within a foreign administrative network domain.
- To provide a reliable network connection between the users, the real-time communication is served via a relay function (TURN Server). This server is deployed in the home domain of the users.
- The overall **delay** for that real-time communication session is **too high** to achieve the best possible QoE (a).
- Based on an external trigger, the TURN Server will be moved with its state from the home domain towards a Mobile Edge Host. A VM Migration is performed.
- While keeping up the TURN Server, the overall delay will be reduced and a better experience for that session will be perceived by the users (b).



RESULTS (BITRATE EVALUATION)

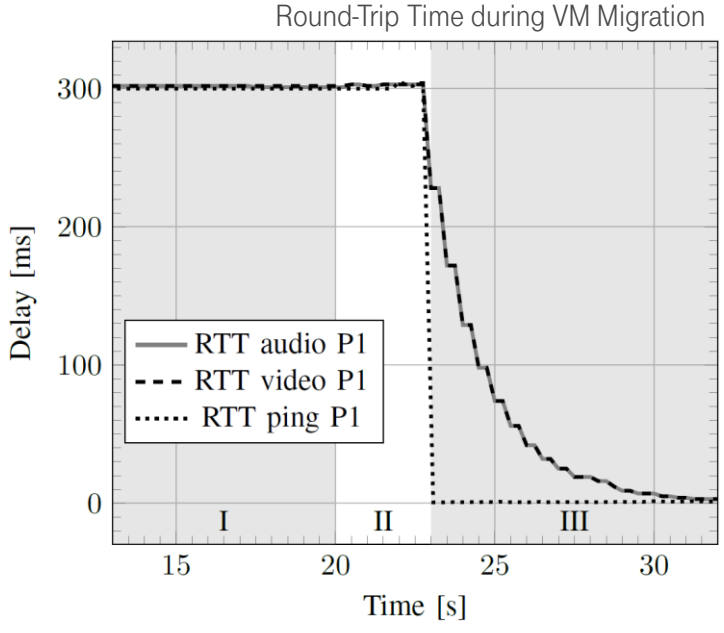
- In the experiments, we measured the received bitrates for audio and video traffic and the packet delay between the both peers.
- Bi-directional audio communication streams of both peers are constant at around 40 - 50 kbps.
- Bi-directional video communication streams have a bitrate of around 2 - 3 Mbps, which is typical for WebRTC video traffic.
- Bitrates vary significantly in Phase II, where the migration process is ongoing.
- The WebRTC engine recognises that something has changed in the network. However, the streams recover really fast to the original values after the migration, as depicted in Phase III.





RESULTS (DELAY/ROUND-TRIP TIME EVALUATION)

- In Phase I, the RTT remains constant at the expected value of ~300 ms (Based on the configured delay of 150 ms for each direction).
- After the migration in Phase II has finished, the value for RTT decreases significantly.
- A smooth drop was recorded with the statistics. The Chrome browser is smoothing down the RTT drop curve progression towards the actual network delay (~ 2 ms) by calculating an average out of former RTT values.
- Parallel ICMP echo requests (Ping) showed the significant drop, as expected when the migration is performed.



CONCLUSIONS AND FUTURE SCOPE

Conclusions

- MEC and their concepts have a high potential for upcoming 5G low-latency applications.
- Our test setup has demonstrated, that the movement of active operating network service functions is technically feasible. Such Live Migration was successfully tested within the presented paper.
- Wrt. the presented use case:
  - The continuity of a real-time communication session retains. However, due to a temporary reduction(1,5 s for audio and 2,5 s for video) of the bit rates we perceived minimal quality degradation in that interval.

Future Scope

- Current evaluations on running Docker containers and session state transfer (Extract/Inject) successfully performed.
- Interfacing towards Application/Service Layer to optimize Service and Network.
- On-Demand Use Case (on Session Setup), the core network components provide the best location of such a Control Point.
- Context information optimizes location for the serving unit (e.g., prediction).

THANK YOU


Kay Hänsge
Master of Engineering
Technology Innovation
Address
Deutsche Telekom AG
Winterfeldtstraße 21
10781 Berlin, Germany
Contacts
Phone: +49 30 8353-54269
mail to: kay.haensge@telekom.de