Mobile Edge Fog Computing in 5G Era: Architecture and Implementation

Shubhranshu Singh Yen-Chang Chiu Yi-Hsing Tsai Jen-Shun Yang
Advanced Communication Technology & Standard Development Dept.
Division for Video & Multimedia Communication Technology
Information and Communication Research Laboratories
Industrial Technology Research Institute
{shubhranshu, cyc, alientsai, jsyang}@itri.org.tw

Abstract— *In the recent past there has been much research and development on cloud based computing. While cloud computing provides huge opportunities, it also imposes several challenges. One of the challenges that current data network operators and future 5G network are foreseeing is huge increase in data traffic. To fulfill such exponential data traffic growth, along with other user expectation, requires multiple innovative approaches and re-consideration of the current network design principals. Mobile Edge Fog computing, as proposed in this paper offers huge opportunities to future data network operators, as well as to equipment vendors. Due to its relatively newer concept, still much research is ongoing and detail architecture is still evolving. This paper proposes and discusses a novel Mobile Edge Computing design and architecture, along with real-time implementation details of the proposed solutions. It exploits many benefits of D2D by incorporating D2D functionalities into the proposed relay-gateways. The application running on top of the proposed network shows significant benefits e.g. in terms of delay as well as core-network signalling and data offloading.*

Keywords— Fog Computing, Mobile Edge Computing, MEC, D2D, ETSI

I. INTRODUCTION

Mobile Edge Fog Computing (MEFC) is a Mobile Edge Computing (MEC) technology which provides cloud-computing capabilities within the access network i.e. in close proximity to mobile subscribers. Due to its proximity to mobile subscribers, it offers a service environment with ultralow latency and high-bandwidth as well as direct access to real-time radio network information that can be used by applications and services to offer context-related services. These services are capable of differentiating the mobile broadband experience. MEFC allows content, services and applications to be accelerated, increasing responsiveness from the edge. The mobile subscriber's experience can be enriched through efficient network and service operations, based on insight into the radio and network conditions.

Data network or cellular operators are increasingly facing new challenges with ever growing mobile traffic. The cloud computing specific applications along with the deployment of Internet of Things will further congest networks. Besides, several applications and content providers are challenged with the latency of the network when connecting to the cloud. Enterprises want the ability to engage with their customers with more efficient, secure and low latency connections. Moving into the 5G era, several out-of-the box design and solution need to be incorporated from the beginning in order to satisfy such growth in mobile traffic. MEFC addresses and solves many of these issues. Data network operators can open 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. Proximity, context, agility and speed can be translated into value and can create opportunities for mobile operators as well as service and content providers.

The remainder of this paper is organised as follows: Section II provides ETSI ISG Mobile Edge Computing and D2D communication as part of background work. Section III discusses design and architecture details of the proposed D2D based Mobile Edge Computing. D2D based MEC local cloud, Relay gateway and MEC platform are discussed within section III. Section IV highlights our real time implementations and section V provides some insights into the performance in terms of delay and signalling optimizations. The paper concludes in section VI, with some thoughts on possible future work and direction.

II. RELATED WORK

A. ETSI ISG Mobile Edge Computing

Mobile Edge Computing has started to receive much attention in the recent days. It is seen as a key emerging technology and an important component of future generation networks, including 5G. ETSI published a white paper on Mobile Edge Computing [2]. This white paper outlines how Mobile Edge Computing is characterized by low latency, proximity, high bandwidth and real-time insight into radio network information and location awareness. It provides examples of how applications such as connected vehicles, e-Health, industry automation, augmented reality, intelligent video acceleration, gaming and IoT services can each benefit from Mobile Edge Computing.

ETSI has recently established an Industry Specification Group on Mobile Edge Computing called ISG MEC, to develop a standardized, open environment that will allow efficient and seamless integration of third-party applications across multi-vendor platforms [1]. The ISG MEC will specify elements that are required to enable

978-1-5090-3438-3/16 $31.00 © 2016 IEEE
DOI 10.1109/ICS.2016.150

applications to be hosted in a multi-vendor Mobile Edge Computing environment. It will enable applications and services to be hosted above the network layer.

B. D2D Communication

Device to Device or D2D communication refers to “direct mode” or “locally routed” path for communication between UEs. The concept of D2D has been there since quite some time e.g. WiFi Direct, Bluetooth, etc. 3GPP has been developing specifications related to D2D, also called as Proximity Based Communication or ProSe, as part of its Release-12 and Release-13 work items. The ongoing work is focused on two aspects: ProSe discovery and ProSe communication and are specified in TS 23.303 [5]. In order to support public-safety uses, 3GPP has also specified protocols to enable UE-to-Network Relays, thus enabling out-of-coverage UE to communicate via UE-to-Network Relays. TS 36.211 specifies RAN aspects including synchronization signal design and synchronization procedure, type 1 and type 2b discovery, physical layer design for discovery which includes resource allocation and discovery signal design, mode 1 and mode 2 communication, L3-based UE-to-Network Relay and D2D for inter-frequency and inter-PLMN discovery. These specifications include protocols to support one-to-one as well as one-to-many communication. 3GPP has now started its work on Release-14. Among other aspects, Rel-14 is likely to include work on evolution of ProSe specification as well as new aspects such as V2X.

III. Proposed MEFC Design and Architecture

Our high level MEFC architecture is shown in the below figure 1. It consists of multiple Relay gateways, which on one side connects to the end devices and on the other end connects to the core network. For our in-lab implementation, relay gateways functionalities are included in the WiFi APs. Among others, these relay gateways include D2D capabilities and routing functionalities to assist in locally requested resource search.

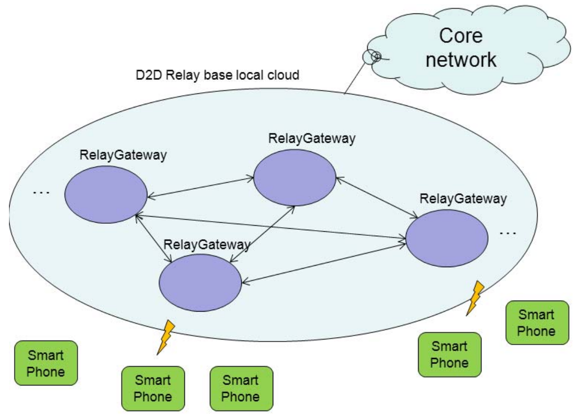

Figure 1. General Design and Architecture of MEFC

Proposed design not only can provide applications or services in a highly decentralized computing environment, but also allows users to complete data processing under mobile communications. Mobile users can experience personal Web APP on MEC platform after registration. Such platform provides high-speed bandwidth and opens up new business opportunities.

A. D2D Based MEC Local Cloud

The relay gateways, as shown in figure 1 provide functionalities of local cloud supporting smart routing and offloading. These relay gateways operate at 5 GHz and implements WiFi direct specifications [15]. It enables communications between devices by connecting D2D relay gateways to form a type of mesh network where each of the relay gateways are connected to each other. Such deployment architecture allows faster deployment, flexible adjustment and network coverage extension. The implemented routing protocol enables intelligent routing thus helping D2D relay network to quickly recover from any link failures.

B. Relay Gateway

As mentioned above, D2D is used to connect relay gateways with each other. Besides, the end devices can connect to the MEC network using their D2D interface. Figure 2, shows details of our proposed relay gateway architecture.

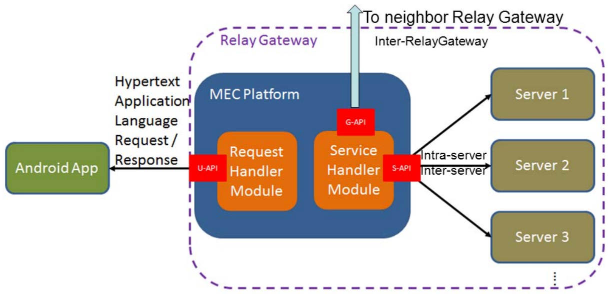

Figure 2. Proposed Relay Gateway Architecture

Each D2D relay gateway includes a MEC platform and may also acts as local server. There are two modules inside MEC platform. These are: request handler module and service handler module. Request handler module handles service request from smart devices. Service handler module brokers to server or to neighbor D2D Relay Gateway in order to load balance among different relay gateways. Figure 3 shows steps involved in request handler module and service handler module. Details of these steps are as below:

Step 1: End device e.g. smart phone sends a service request to the nearest D2D relay gateway requesting access to a specific service.

Step 2: On receiving service request from smart phone, request handler module of relay gateway, which is implemented as part of MEC platform of the connected D2D relay gateway, parses the request to get the service information. Relay gateway then accordingly forwards this request to the service handler.

Step 3: Service handler, based on the received service information, checks if this particular service exist.

Step 4: If the service exists locally, then the service handler processes the service request. Otherwise, it selects a relay gateway as per algorithm shown in Figure 4 and then forwards the request to its service handler for processing this task.

Step 5: Service handler gets a result at the end of service process.

Step 6: After successful completion, service handler returns result to request handler. The request handler encodes the result to a universal format and then returns to smart phone.

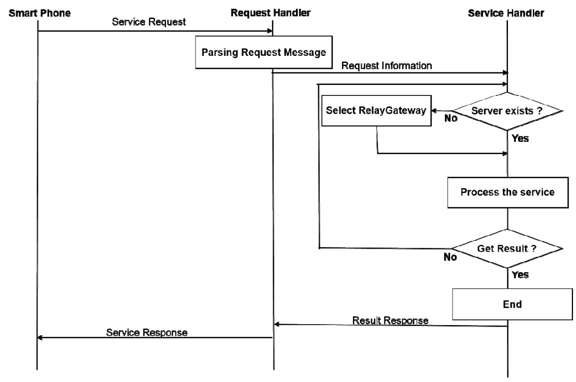

Figure 3. Flowchart of request handler and service handler

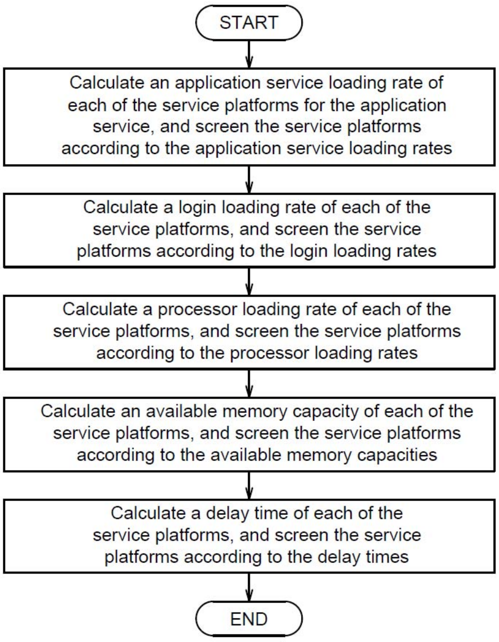

Figure 4. MEC platform selection algorithm

Design and algorithm as explained above enables software developers to integrate different types of services via a single interface. It helps to integrate local cloud services and smart devices in a fast and easy way.

C. *MEC Platform*

Figure 5 shows data flow and related details in scenarios where multiple MEC platforms exist. Dotted lines and solid lines are used to indicate connection relationship between the elements, and the solid lines indicate data transmission in the present example. The MEC platform 1 has application services SR1, SR2, and so on. The MEC platform 3 has application services SR1, SR3, and so on. The MEC platform n has application service SR1, SR4, and so on.

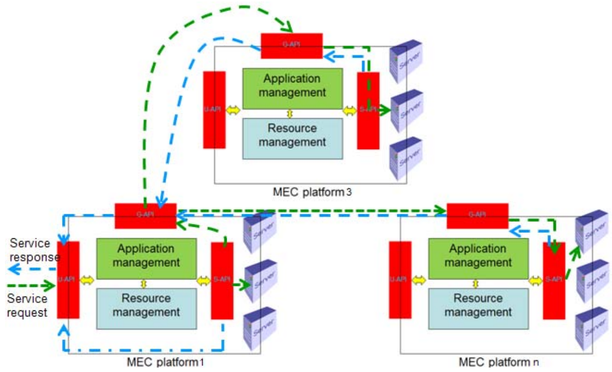

Figure 5. Data transmission diagram

At first, the client transmits a request message to one of the MEC platforms through User API or U-API. In case of our in-lab implementation, request message is written in a standard hypertext application language (HAL). After the U-API of the MEC platform receives the request message, the resource management analyzes the content of the request message and virtualization of hardware and software. According to the application service of the request message, the application management provides the middleware services of the applications which are hosted on MEC platform. The Server API (S-API) controls the communication of application and server registry.

Figure 6 shows flowchart of a control method for MEC, especially in case of Inter-MEC platform communication. As shown in Figure 6, S-API determines whether the request message is for another application service or not. If the request message is to application service in local MEC platform, then the method terminates. On the other hand if it is meant to be for another application service not locally available, then it selects one of the service platforms according to first application service of the request message.

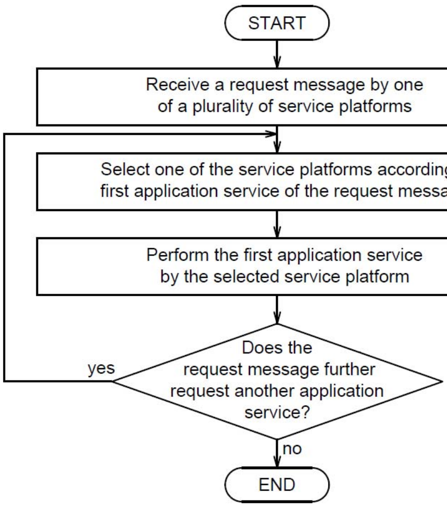

Figure 6. Flowchart of a control method for MEC

In general, unified Intra/inter Service API Protocol for Service Handler function is to:

- Unify APIs between services to service, service to host and host to host

- Manage Edge Services

- Increase Software development and integration efficiency, and facilitate easy functional integration

Relay gateway API (G-API) updates its hardware capability dynamically which include CPU usage, available memory capacity, etc. To achieve short latency and minimal delay time in D2D relay network, G-API that is relay gateway Common API may provide integration of Inter/Intra Relay Gateway connectivity.

D. IMPLEMENTATION

Our Mobile Edge Fog Computing implementation, with supported applications is as shown in Figure 7. Following is D2D relay based MEFC network application scenario:

Smart Devices receives Discovery Announcing Message from Edge Cloud. In our case, the smart devices used are ITRI glasses, which can also be replaced by Google Glasses.

Smart Devices or ITRI Glasses send enquiry request to NMS (Network Management System) for the nearest Server Location via LTE connection.

NMS sends to the nearest Edge Cloud Server IP address or any other ID to ITRI glasses based on received information, NMS requests specific D2D relay GW to serve ITRI glasses.

ITRI glasses send captured Logo image to specific D2D relay GW. Gateway performs Parallel Process Technology to execute image recognition process.

Edge Cloud Server identifies the Logo image and sends promotion video/e-coupon to D2D relay GW and push to phone

Edge Cloud Server displays promotion video on nearby LCD to enrich shopping experience

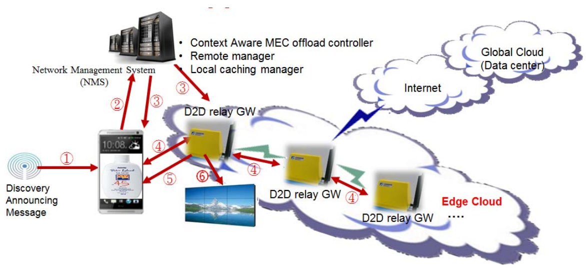

Figure 7. D2D Relay Based MEFC Network Application Scenario

IV. PERFORMANCE EVALUATION

Figure 8 shows the average execution times of download for each file over 10 times on both local cloud and, for the comparision purposes, from google drive located on the Internet. In the local cloud, the data stores on local server and the client connects directly to it. In the case of google drive, data are stored on google drive cloud space located on the Internet. As can be seen from figure 8, the transmission time of local cloud access is less than google drive access. To get a file from a local cloud cache is much more reasonable than it is from cloud.

As edge services run close to end devices it considerably reduces latency. Table 1 shows the packet round-trip time between end devices and server. Since end devices connected directly to local cloud, it only goes through 1 hop and a few round-trip times. On the other side, end devices trying to fetch data from google drive go through many routers and the round-trip time is much longer. Thus accessing services from edge can be utilized to significantly reduce latency, to improve user experience, or to minimize congestion in other parts of the network.

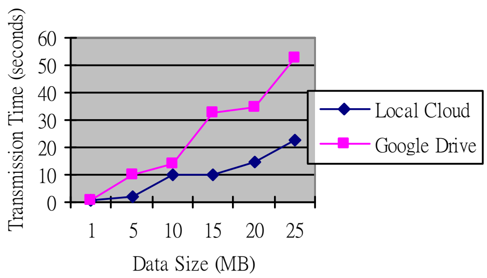

Figure 8. Transmission Time versus data size

Table 1. Average Hop Count and Average Round-trip Time

	Hop	Average RTT
Local Cloud	1	1 ms
Google Drive	13	61 ms

A. Conclusion and future work

This paper provides real-time implementation details of a D2D based Mobile Edge Fog Computing system. This is based on the proposed novel and enhanced MEC architecture which includes D2D based MEFC local cloud, relay gateway and MEC platform. The performance from the implemented real-time system highlights several of the benefits both from the end user perspective as well as from network perspectives. The proposed solution minimizes round trip time and maximizes throughput for optimum quality of experience. It can be used to offer consumer or enterprise propositions, such as tourist information, sporting event information, advertisements etc.

A future work can be to further provide solutions to several challenges of distributed caching in the proposed MEFC architecture. A distributed caching architecture provides backhaul and transport savings and improved QoE. Local content caching has the potential to reduce backhaul capacity requirements to an estimated vaue of up to 35%.

Acknowledgment

This research is supported by the Information and Communications Research Lab (ICL), Industrial Technology Research Institute (ITRI), Taiwan via project Broadband Convergence System and Integration Technologies funded by MOEA.

References

[1] ETSI Mobile Edge Conputing: http://www.etsi.org/technologies-clusters/technologies/mobile-edge-computing

[2] White Paper: ETSI's Mobile Edge Computing initiative explained https://portal.etsi.org/Portals/0/TBpages/MEC/Docs/Mobile-edge_Computing_-_Introductory_Technical_White_Paper_V1%2018-09-14.pdf

[3] Mobile-Edge Computing (MEC); Service Scenarios: ETSI Standard Number: GS MEC-IEG 004 http://www.etsi.org/deliver/etsi_gs/MEC-IEG/001_099/004/01.01.01_60/gs_MEC-IEG004v010101p.pdf

[4] Mobile-Edge Computing (MEC); Proof of Concept Framework: ETSI Standard: GS MEC-IEG 005 http://www.etsi.org/deliver/etsi_gs/MEC-IEG/001_099/005/01.01.01_60/gs_MEC-IEG005v010101p.pdf

[5] 3GPP SA2 TS: 23:303: Technical Specification Group Services and System Aspects;Proximity-based services (ProSe); Stage 2

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[8] "Mobile-Edge Computing – Introductory Technical White Paper," ETSI, Sep. 2014, URL: http://www.etsi.org/

[9] " Executive Briefing – Mobile Edge Computing (MEC) Initiative," ETSI, Sep. 2014, URL: http://portal.etsi.org

[10] Zhongren Cao , Matthew French , Rajesh Krishnan , Joshua Ng , David Talmage and Qinqing Zhang , "Content-Oriented Mobile Edge Technology System Integration Framework and Field Evaluation," 2014 IEEE Military Communications Conference, pp. 1405-1410, Oct., 2014.

[11] Hsien-Wen Chang, Kun-Yi Lin and Yen-Chang Chiu, "A Multihop Routing D2D Relay Gateway and Networking Technology," ITRI Journal of Information and Communication Technology No.160, pp.53-59, Dec. 2014.

[12] Hsien-Wen Chang, et al., "A Proximity-Based Message-Relaying Mechanism Using Wi-Fi Direct Technology for Public Safety Systems," 2014 IEEE 11th Asia Pacific Wireless Communications Symposium （APWCS 2014）, 28-29 Aug., 2014.

[13] Michael Fagan , Mohammad Maifi Hasan Khan and Bing Wang , "Leveraging Cloud Infrastructure for Troubleshooting Edge Computing Systems," 2012 IEEE 18th International Conference on Parallel and Distributed Systems, pp. 441-447, Dec., 2012.

[14] "Boosting User Experience by Innovating at the Mobile Network Edge," ETSI, Oct., 2014, URL: http://www.etsi.org/

[15] WiFi Direct: http://www.wi-fi.org/discover-wi-fi/wi-fi-direct