Agile management of 5G core network based on SDN/NFV technology

Taesang Choi, TaeYeon Kim, Wouter TaverNier, Aki KORVALA, Jussi PAJUNPAA

Electronics and Telecommunications Research Institute (ETRI) Daejeon, Republic of Korea
Email:{choits, tykim}@etri.re.kr

GHENT Univ., Belgium
Email:Wouter.Tavernier@UGent.be

Nokia, Finland
Email:{aki.p.korvala, jussi.pajunpaa}@nokia.com

Abstract— In the 5G era, radio IP capacity is expected to reach 20 Gbit/s per sector, and ultra-large content traffic will travel across the faster wireless/wireline access network and packet core network. Also massive and mission-critical IoT is the main differentiator of 5G services. These types of real-time and large-bandwidth consuming services require radio latency of less than 1ms, and end-to-end latency of less than a few ms. By distributing 5G core nodes closer to cell sites, backhaul traffic volume and latency can be significantly reduced by having mobile devices downloading content immediately from a closer content server. In this paper, we propose a novel solution based on SDN (Software Defined Network) and NFV (Network Function Virtualization) technology in order to achieve an agile management of distributed 5G core network functionalities and services with PoC implementation targeted for Pyungchang Winter Olympics.

Key words: 5G core, agile management, Software-Defined Network (SDN), Network Function Virtualization (NFV);

I. Introduction

In the 5G era, radio IP capacity is expected to reach 20 Gbit/s per sector (mobile speeds up to 20 Gbit/s), and ultra-large content traffic (e.g. UHD Video Streaming, Augmented Reality (AR), Virtual Reality (VR)) will travel across faster wireless/wireline access network. All 5G mobile/fixed traffic has to travel via packet core network. Currently, in 4G, most mobile operators (even large-scale ones) have only a few sites with PGWs across their entire networks.

The SDN paradigm provides a new capability for faster service provisioning of the 5G core network through standard programmable interfaces. Also with the cloud computing, datacenters promote on demand provisioning of computing resources and services [1]. If 5G core nodes are distributed closer to cell sites, content servers (or caching servers) can be placed on the rack right next to the distributed 5G core with NFV technologies. This can help significantly reducing backhaul traffic by having mobile devices downloading content immediately from the content server. Thus it is desirable to distribute their packet core functionality to a number of local sites near to the end users in the coming 5G era. 5G core functionality and applications can then run on virtualized servers at the local network sites. Another important 5G services, massive and mission-critical IoT services, are the main differentiator against 4G services. Mission-critical IoT (Ultra-reliable and low latency communications) applications include remote controlled machine, autonomous driving, etc. These types of ultra-real-time services require radio latency of less than 1ms, and end-to-end latency of less than a few ms [2].

To address such challenges, we propose an intelligent 5G management and orchestration system enabling agile management of 5G distributed core network functionality and services.

In this paper, we present key technologies, management and orchestration architecture, our implementation, and deployment experiences with proof of concept use cases targeted for Pyungchang Winter Olympics. The proposed solution is an interim result of collaboration project between Korea (KR) & EU. The rest of the paper is organized as follows. We describe the design principles in Section II. We present our management and orchestration system architecture in Section III. Our prototype implementation and deployment experiences are described in Section IV. Finally, we conclude our effort with potential future work in Section V.

II. Key Technologies and Architecture

This section examines key technologies on SDN, NFV, Management and Orchestration (MANO), Mobile Edge Computing (MEC), and mobility management and their associated architectures for the support of the proposed agile management of core network functionality and services.

- **Software Defined Networking and Orchestration**

Standardization efforts on SDN were made mainly by ONF[3] and ITU-T SG13[4] by defining requirements, reference architecture, protocols, and use cases. Open source projects such as ODL (Open Daylight)[5] and ONOS (Open Networking OS)[6] played an important role to realize SDN concept in real life. It started with limited networking environment such as cloud data centers and enterprise networks and widen its coverage into wide area transport network and wireless/wireline integrated multi-domain networks. Instead of applying it as a standalone network control tool, it is now used with NFV and as a component of an end-to-end orchestration solution. It provides an intelligent knowledge plane to make control decisions via traffic steering, traffic engineering, and flexible service chaining for latency sensitive and reliability seeking applications. It can be used in efficient communications among distributed core functional components.

- **Network Function Virtualization**

Virtualization of core and radio access network functions will optimize the use of network resources, add scalability and agility. To this end, the ETSI NFV ISG (Industry Specification Group) has defined the architecture, open APIs and reference points, leveraging open source proof of concept (PoC) projects and communities to drive open standards of NFV. In 2016, it published Release 2 specifications and reports, including functional requirements, interface and information model of reference points for the management and orchestration function block, called NFV-MANO[7]. These open standards are intended to enable third party vendors to develop framework components that can collaborate with various vendor components so that CSPs are not restricted in selecting functional and management components.

The main appeal of using NFV to deploy network elements and virtual network functions (VNFs) is that services can be launched more quickly, by installing software on a standard hardware platform. This is akin to the way software applications could be developed and launched for the PC platform when it first emerged. Another advantage is lower capital expenditures, because standardized hardware platforms tend to drive cost down. Such advantages can be directly applied to the distributed core functional components communications environment.

- **Mobile Edge Computing**

In order to support requirements on market's expected throughput, latency, scalability and programmability, ETSI established an Industry Specification Group on Mobile Edge Computing in 2014. It develops a standardized and open environment that offers distributed cloud-computing capabilities and an IT service environment to application developers and content providers. By February 2016, the group has finalized three specifications: Terminology, Technical Requirements and the Framework and Reference Architecture. It also works on specifications for MEC platform Application Enablement, API principles and guidelines, Services APIs for Radio Network Information and Location, UE identity and Bandwidth management, system/host/platform management, lifecycle and policy management, UE application interface, Deployment of MEC in a NFV environment, and End-to-end Mobility

By offering distributed cloud-computing capabilities and exposure to real-time radio network and context information, it provides the following characteristics:

- Ultra-low latency: Mobile Edge services can be run close to the end user devices to provide the lowest possible latency;
- Proximity: Being close to the source of information, Mobile Edge Computing is particularly useful to capture key information for analytics and big data;
- High Bandwidth: Mobile Edge location at the edge of the network combined with the use of real time radio network information can be used to optimize the bandwidth for the applications;
- Location awareness: Mobile Edge can leverage the low-level signaling information to determine the location of each connected device;
- Real time insight into radio network and context information: Real-time network data can be used by the applications and services to offer context-related services.

It can bring significant improvement of mobile user's Quality of Experience on latency or QoS sensitive services such as Edge Video Orchestration, Mobile Video Throughput guidance, Augmented reality, Intelligent Video Analytics etc. Most importantly, MEC enables the implementation of mobile edge applications as software-only entities that run on top of a virtualization infrastructure, which is located in or close to the network edge.

- **Distributed Mobility Management**

It is essential to support distributed mobility management to enable agile management of core network functionality. Currently IETF is conducting standardization efforts on defining distributed mobility management architecture and mechanism on layer 3 IP network environment. 3GPP also initiated work on defining layer 2 distributed mobility management requirements for mobile communications environment. Functional decomposition & distribution for global service management will span multiple PoP(Point of Presence)s over the network including network slices in 5G environment. Anchoring and mobility management tailored to such a network environment better to be determined at the central node unlike exiting hierarchical and IP mobility. Composition functions and resource will be orchestrated for dynamic mobility management. Various experiments & simulations are under study by research community and extensive testing and verification of the concept of the distributed mobility management are needed.

- **Core Network Architecture**

We designed our core network architecture (Figure 1)[8] to support agile management of core network (CN) functionality based on various key technologies described in this section. Specifically, the CN functionality is realized by leveraging SDN and NFV in order to facilitate the dynamic provisioning of CN functions. By using SDN capabilities, dynamic control of traffic flows can be performed, redirecting the traffic to gateways according to, say, workloads. Simultaneously, the introduction of NFV permits the separation of service functionalities from the capacity constrained specific network entities and allow dynamic instantiation in commodity and powerful servers. Extensions to OpenFlow are being developed which would help create a SDN/NFV-based Mobile Packet Core [9].

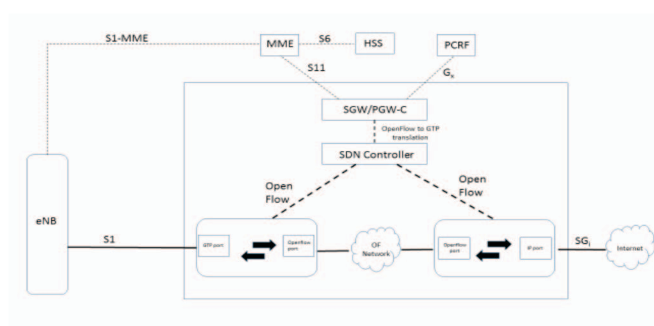

Figure 1. Core Network Architecture

Main core network functions are designed and implemented in the form of virtual functions, namely, vEPC (virtual evolved packet core). Both EU and KR sides provide their own implementation of vEPCs based on this architecture. They are described as follows;

- **European vEPC architecture (5GTN)**

The EU vEPC consists of the following VNFs:

- Mobile gateway: The Cloud Mobile Gateway (CMG) provides the SP-GW, GGSN and Traffic Detention Functions (TDFs), evolved packet data gateway (ePDG) and trusted wireless access gateway (TWAG).
- Mobility management: The Cloud Mobility Manager (CMM) provides the mobility management entity (MME), and SGSN functions.
- Policy control and charging: The Dynamic Services Controller (DSC), built on the patented Agile Rules Technology (A.R.T) rules engine, provides the Policy and Charging Rules Function (PCRF) and wireline Radius/Change of Authorization.
- Element and network management: The Service Aware Manager (SAM) provides end-to-end network management visibility across the entire mobile network.

To support the scalability required to meet the expected 5G and IoT service requirements, the Packet Core VNFs provide three key design innovations:

- Packet core VNFs are decomposed into separate control and data plane virtual machine (VM) instances. This enables a distributed architecture where data plane resources can be deployed in edge data centres, closer to the device while control plane resources can be centralized.
- State-efficient VNF processing, which unpins the subscriber/device state information from the VMs, freeing up the underlying compute resources to be reused to process other subscribers/devices.
- Remote cloud database, which syncronizes the subscriber/device state information into a real-time data store.

The 5GTN functional architecture[8] is given in Figure 2.

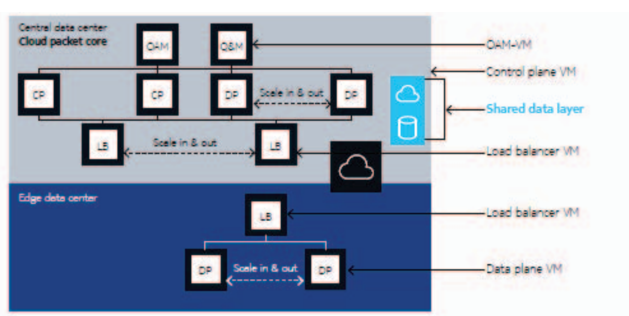

Figure 2. 5GTN vEPC functional architecture

- **Korean vEPC architecture**

The core network of 4G LTE is in charge of mobility, authentication and charging allowing all mobile traffic pass the core network to access services incurring traffic congestion in the core networks. Our architectural decision for 5G is to distribute mobile core functions to the edge nodes. 5G core is generally divided into 5G Core UP (User Plane) in charge of bearer delivery and 5G Core CP (Control Plane) in charge of signalling and control of the 5G core network. The key CN architectural design principle is a centralized CP with distributed UP over the edge nodes.

If the core network where bearers are terminated located closer to cell sites, application servers follow naturally and backhaul traffic will significantly decrease, bringing in cost reduction for continual backhaul enhancement.

5G network is supposed to be able to provide ultra-real time services like highly sensitive remote control, automatic driving vehicle, etc. These types of services may generate much lesser traffic than video streaming applications, but require ultra-low latency. Figure 3 illustrates the high-level architecture of Korean vEPC. It is realized as HSvEPC (Highly Scalable virtual Evolved Packet Core)[8]. Its functionality and architecture are described below.

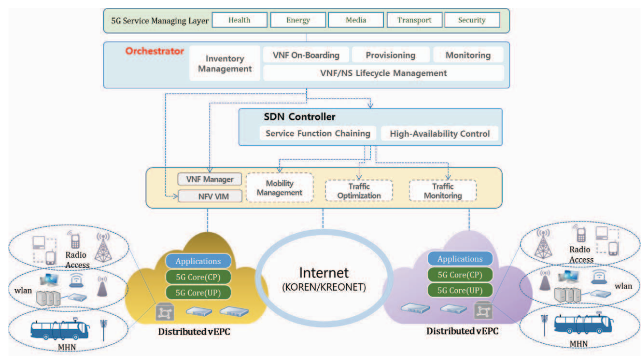

Figure 3. High level architecture of Korea's Distributed virtualized EPC

- **HSvEPC Network Architecture**

It is possible to deploy different types of virtual mobile packet core depending on demand or network access environment in HSvEPC network architecture. Two types of vEPC are desiged:

- S-vEPC (splitted vEPC): The first type is an expansion of vEPC by separating conventional consolidated functions into user plane and control plane functions for dynamic scaling operations.
- MHN-vEPC: Another type is an optimized case for hot spot area to enhance agility of the network. For faster and more dynamic mobility management in the Mobile Hot Spot area, S1 interface of the virtual EPC has been modified in terms of user plane and control plane.

Figure 4 and 5 shows the HSvEPC functional architecture.

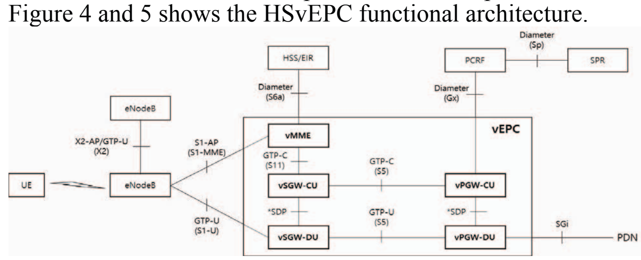

Figure 4. HSvEPC Functional Architecture - S-vEPC

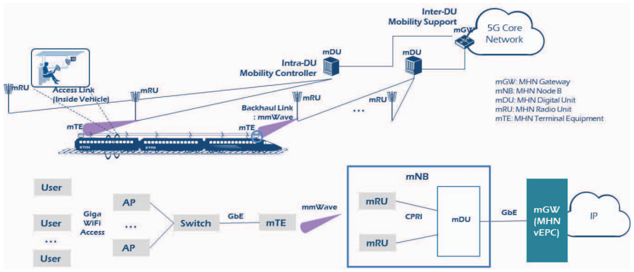

Figure 5. HSvEPC Functional Architecture - MHN-vEPC

- **Management and Orchestration Architecture**

Figure 6 shows EU side overall management and orchestration architecture based on NFV MANO and SDN.

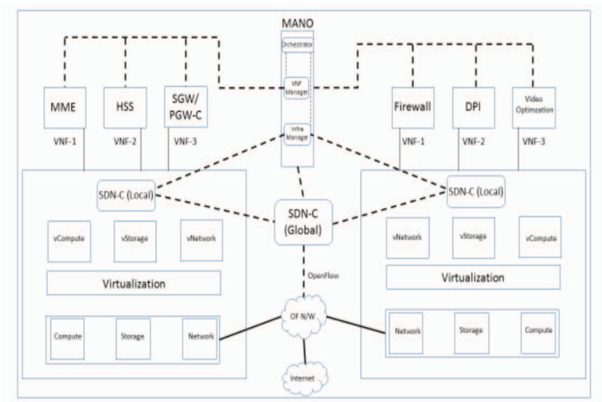

Figure 6. Overall M&O architecture with MANO & SDN

The architecture has two management entities:

- VNF manager – is in charge of instantiating and controlling the EPC functions. It is responsible for interacting with the VNFs, chaining of the VNFs and handling their lifecycle – instantiation, maintenance etc. It is in-charge of the operation and configuration of VNFs, through the operations support system (OSS)/base station subsystem (BSS). It will handle multi-functional EPC components like the MME, HSS etc. as well as the specific functionality VNFs like Firewalls, Deep Packet Inspectors etc.
- Infrastructure manager – This entity interacts with (or incorporates the capability of) the SDN controller in the service stratum when deploying the VNFs for configuring the computing and storage resources for the VNF of interest. It also supports for the networking part to attach the VNFs to the border of the underlying transport network to make them reachable from outside the data centre. This is only for the service layer part. It also has to decide a path for the transport layer VNFs.

KR core network management and orchestration are also based on MANO and SDN. Figure 7 shows the management and orchestration architecture.

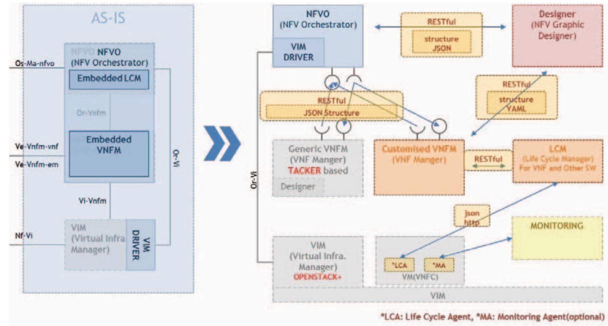

Figure 7. KR M&O Architecture

It is comprised of three different entities — NFV Orchestrator, VNF Manager and Virtual Infrastructure Manager (VIM).

NFV Orchestrator is responsible for managing the functions such as network service life-cycle management and the overall resource management. Service management or orchestration deals with the creation and end-to-end management of the services by composing different VNFs. Resource management helps ensuring that the NFV-infrastructure resources are abstracted cleanly (independent of VIM) to support the services that access these resources.

The VNF Manager oversees the lifecycle (typically involves provisioning, scaling, terminating) management of instances of VNF. In this case, each VNF is associated with a VNFM that will manage that particular VNF's lifecycle. A VNFM may manage multiple instances of the same type of VNF or different types of VNFs.

The VIM controls and manages the NFVI compute, storage, and network resources. The VIM-component has received tremendous focus and various open source solutions such as OpenStack and has been used to realize the virtualized infrastructure management functionality of MANO.

- Auto scaling based on Performance/Fault Management

In our M&O, auto-scaling functionality is provided as shown in Figure 8. After instantiation of the 5G mobile core network service, NFVO sends supervision request to the supervisor, which performs performance monitoring, and fault notification over virtualized resource and functions. Scaling is conducted autonomously by the orchestrator based on the information notified from the supervisor.

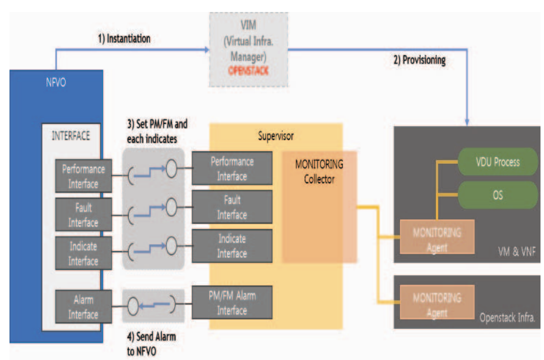

Figure 8. MANO Auto-scaling Process

- Automation by event chaining

Another important functionality supported is an event chaining process, which is defined as the sequence of event units occurring from inside or outside of the target VNF and VDU. It enables automation of the 5G mobile core network management. A combination of internal events significant in a single VNF or VDU and external events between VNFs and VDUs enables automated management of lifecycle of mobile core network service.

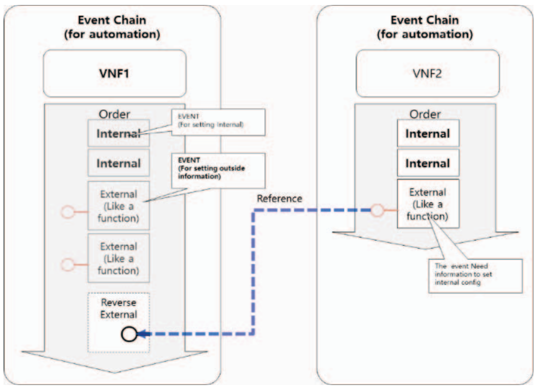

Figure 9. MANO Automation Process

III. IMPLEMENTATION AND DEPLOYMENT EXPERINECE

Both EU and KR edge and core network functions are under development. Development of some components has been completed such as EU's edge and core functions under 5GTN solution. KR's vEPC development is underway with core functionality completed. KR vEPC currently support up to 100 UE and 20 Gbps channel throughput toward eNB. To meet the 5G KPI, we are trying to fill the gaps in both vEPC systems. We are targeting to complete our system development by June of 2017.

We are also developing our agile core network management and orchestration systems based on the architectures described in the section II. Initial prototypes are available and are testing functionality as separate systems and interoperability as well across EU and Korea over interconnected R&D networks between EU and KR via KOREN-TEIN-NORDUNET-FUNET.

- **vEPC Implementation**

5GTN vEPC VNF functions have been implemented, deployed and tested on CloudBand's NFV infrastructure (NFVI) and its management and orchestration (MANO) solution. CloudBand is a hardened, production-ready NFV solution based on OpenStack and other open source technologies. This open approach allows service providers to benefit from a vast community of engineers and supports investments in a mainstream solution with open interfaces.

HSvEPC implementation consists of vMME, vSGW-CU, vPGW-CU, vSGW-DU, and vPGW-DU. CU is a control plane which controls device management and DU is a user plane which controls data transfer between devices. The main reason why we separated functions by each plane is to support scalability depending on demand situation. Since the functions in HSvEPC are implemented as VNFs, they can be modified on demand and controlled per VNF level. One important use case of such flexibility is network slicing support.

The figure 10 describes APN (Access Point Name)-based core network slicing use case. An IoT device may have a different APN against an UE and discrimination of each device at MME is required. The above use case illustrates our implementation of MME that can classify different devices by categorizing based on their APNs and map into appropriate resources in SGW and PGW. Also HS-vEPC can be scaled in or out depending on demand which can reduce cost and other unused part of network function and relocated into the only necessary parts.

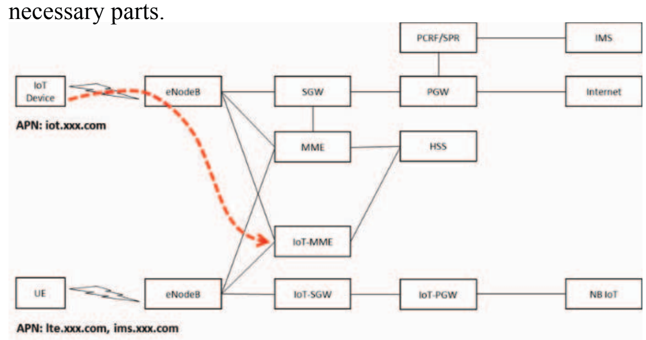

Figure 10. HSvEPC Core Slicing Use Case

- **Management and Orchestration Implementation**

MANO in 5GTN has been implemented and deployed. It consists of CloudBand Infrastructure Software, CloudBand Application Manager, and CloudBand Network Director that have been optimized to fit the key NFV management and orchestration (MANO) shown in figure 11.

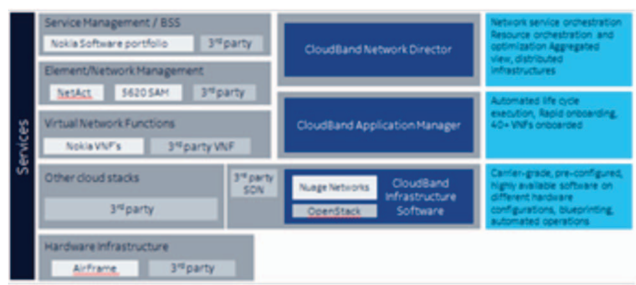

Figure 11. 5GTN MANO Implementation

- CloudBand Infrastructure Software
  CloudBand Infrastructure Software is a multi-purpose NFV infrastructure (NFVI) and infrastructure manager (VIM). It virtualizes and manages compute, storage, and network resources.
- CloudBand Application Manager
  CloudBand Application Manager is a VNF manager (VNFM) that automates lifecycle management actions by managing resources and applying associated workflows.
- CloudBand Network Director
  CloudBand Network Director is an NFV resource and network service orchestrator. It manages virtual resources across geo-distributed NFV infrastructure nodes. It visualizes and automates the lifecycle of network services, such as virtual CPE, including their forwarding graphs and service chains.

KR MANO implementation is shown in Figure 12. We have implemented it in a rack of servers consisting of VIM built and extended over OpenStack, a VNF manger, and an orchestrator. The management target is, of course, a set of virtual functions implementing core network functionality and networks that interconnects those virtual core functions.

Figure 12. KR MANO Implementation

- **Deployment and EU-KR interoperability Testing**

Field conformance and interoperability testing between EU and KR will be held during July to September timeframe[10]. We defined two scenarios in which the EU-KR 5G interoperability has to be demonstrated:

- Scenario 1 – Where there are two users, one connected to the EU EPC and the other to the KR EPC. Content is shared between the two users which is a latency critical application like shared gaming.
- Scenario 2 – In this scenario a mobile UE on the KR side is the content provider and is streaming UHD videos to a receiving UE on the EU side. The aim is to achieve very high data rates across the two EPCs.

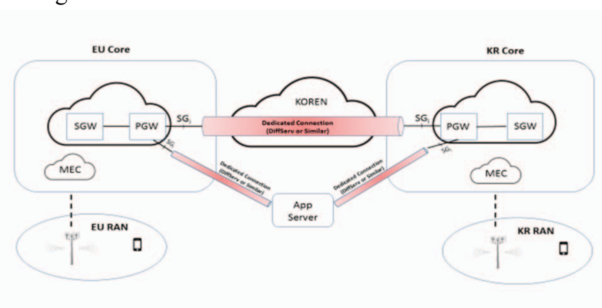

Figure 13. EU-KR Loose Interoperability Testing Architecture

For the first phase, we are going to conduct a loose interoperability testing (Figure 13) as defined below:

- Standard SGi-like PDN interconnection via IP.
- Dedicated tunnel between EU and KR test-bed which will provide guaranteed bandwidth and latency – This will ensure that the QoS requirements of the two use-cases are guaranteed.
- Model similar to the DiffServ model can be used here to guarantee QoS. But this model must be capable of providing 5G standard QoS. Details of such a model needs to be worked out further.
- Reachability fixed IP or DNS based system, depends on the actual applications for both use case scenarios defined.
- Support for dual stack – both IPv4 & IPv6 will be supported.
- Dynamic routing protocols (OSPF, BGP) for advertising PGW IP address to external network
- Application server placed strategically between the two cores which will enable the running of common applications like gaming etc. with low latencies. The connection to and from these servers will also have guaranteed QoS.
-

IV. Conclusion and Future Work

In this paper, we proposed an intelligent 5G management and orchestration system enabling agile management of 5G distributed core network functionality and services to address 5G KPIs. As details of the proposed system, we presented key technologies, management and orchestration architecture, and our implementation and deployment experiences with proof of concept use cases. As described, we are currently in the phase of conformance and interoperability testing of the proposed system functionality. Our future work includes performance evaluation of the proposed solution in proof-of-concept testing environment. We will incorporate these analysis results in our future version of paper for the completeness.

Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No.B0115-16-0001, 5G Communication with a Heterogeneous, Agile Mobile network in the PyeongChang wInter Olympic competioN) and European Union H2020 5GPPP under grant n. 723247

References

[1] H. Shokri,-et.All, "Millimeter Wave Cellular Networks: A MAC Layer Perspective," in IEEE Transactions on Communications, vol. 63, no. 10, pp. 3437-3458, Oct. 2015.

[2] V. Desai, et.All, "Initial beamforming for mmWave communications," 2014 48th Asilomar Conference on Signals, Systems and Computers, 2014, pp. 1926-1930.

[3] Open Networking Forum. http://www.opennetworking.org/.

[4] ITU-T SG13. http://www.itu.int/.

[5] ODL open source project. http://www.odl.org/.

[6] ONOS open source project. http://www.onos.org/.

[7] ETSI Network Functions Virtualisation (NFV) – Management and Orchestration,http://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_nfv-man001v010101p.pdf

[8] Wouter, et. al., 5G CHAMPION architecture, API- and interface document (D2.1), 5G CHAMPION project, Oct. 2016.

[9] ONF, Wireless & Mobile Working Group, (WMWG), https://www.opennetworking.org/index.php?option=com_content&view=article&id=1179&Itemid=463

[10] M. Mueck, *et al.*, “5G CHAMPION - Rolling out 5G at 2018 Winter Olympic Games,” in *Proc. IEEE Global Conf. on Commun.*(*GLOBECOM*), 2016.