Taleb, Tarik; Mada, Badr; Corici, Marius Iulian; Nakao, Akihiro; Flinck, Hannu

PERMIT

Published in:
IEEE Communications Magazine

DOI:
10.1109/MCOM.2017.1600947

Published: 01/05/2017

Document Version
Peer reviewed version

Please cite the original version:
Taleb, T., Mada, B., Corici, M. I., Nakao, A., & Flinck, H. (2017). PERMIT: Network Slicing for Personalized 5G Mobile Telecommunications. *IEEE Communications Magazine*, *55*(5), 88-93. [7926922]. https://doi.org/10.1109/MCOM.2017.1600947

PERMIT: Network Slicing for Personalized 5G Mobile Telecommunications

Tarik Taleb
Aalto University
Espoo, Finland

Marius-Iulian Corici
Fraunhofer Fokus
Berlin, Germany

Akihiro Nakao
University of Tokyo
Tokyo, Japan

Hannu Flinck
Nokia Bell Labs
Espoo, Finland

Emails: tarik.taleb@aalto.fi; marius-iulian.corici@fokus.fraunhofer.de; nakao@iii.u-tokyo.ac.jp; hannu.flinck@nokia-bell-labs.com

***Abstract* – 5G mobile systems are expected to meet different strict requirements beyond the traditional operator use cases. Effectively, to accommodate needs of new industry segments such as health care or manufacturing, 5G systems need to accommodate elasticity, flexibility, dynamicity, scalability, manageability, agility and customization along with different levels of service delivery parameters according with the service requirements. This is currently possible only by running the networks on top of the same infrastructure, technology named network function virtualization, through this sharing the development and infrastructure costs between the different networks.**

In this paper, we showcase the need for the deep customization of mobile networks at different granularity levels: per network, per application, per group of users, per individual users and even per data of users. The paper also assesses the potential of network slicing to provide the appropriate customization and highlights the technology challenges. Finally, a high level architectural solution is proposed addressing a massive multi-slice environment.

I. INTRODUCTION

Mobile networks are nowadays architected to serve all mobile users; ensuring some degree of service-level differentiation, by making decisions on different user profiles, but with no specific tailoring of the functioning to the specific user needs. However, statistics demonstrate that users do not behave all in the same way: 53% are light mobile phone users; 24% exhibit medium usage behavior, and the remaining 23% are heavy mobile phone users [1]. Even amongst heavy mobile users, usage patterns of data-intensive mobile applications, namely those related to social, news, and video vary considerably [2][13][14]. From these statistics and others, it becomes apparent that having the same mobile network architecture serving all mobile users, let alone all mobile applications, despite the diversity they exhibit in their attitudinal response to mobile services, have to be rethought[15].

Furthermore, a mobile user usually subscribes to a single mobile operator that provides the delivery of all the mobile services. In addition, a single mobile network usually ensures the communication for all service types, regardless of the suitability of its available functionality to deliver these services with acceptable Quality of Experience (QoE) and network efficiency. Due to the uniformity of the network, all users of the network are charged based on the same monthly subscription model which fails to capture the specifics of the heavy phone users and makes too expensive the network for individual large scale sensor deployments.

Last but not least, current mobile core networks are offering a uniform ubiquitous service for all the connected devices. Even if a mobile user moves far away from the mobile core network infrastructure, he remains serviced by the same core network, even in case of highly decentralized mobile system [3]. This feature may impact numerous emerging advanced mobile services with strict latency and jitter requirements (e.g., augmented reality, online mobile gaming, and location-based social networking). High latency and jitter degrade such mobile services, rendering their respective devices (e.g., Microsoft HoloLens and Google Glasses) unusable, ultimately turning users away and impacting revenues. Also, a large number of users do not move from a specific network area, their mobility support being only an additional overhead on the network.

To cope with the above, this paper advocates the need for personalizing and customizing mobile telco services, through providing a functional differentiation for the different user requirements. Hereby, the intention is to leverage the emerging technologies in the areas of Network Function Virtualization (NFV), Software Defined Networking (SDN), and cloud and edge computing for providing a single infrastructure on top of which multiple versions of the same software, generically named slices, customized to have specific behaviors are running, through this removing the overhead of the uniform network service.

Furthermore, in order to account for different layers of granularity, this paper underlines the need and the technical possibilities to support a very large number of slices as well as their appropriate network placement according to the momentary location of the subscribers.

The paper is organized as follows: Section II presents quick overview on network slicing and highlights its utility for the PERMIT (**Per**sonalized **M**ob***i***le **T**elecom) vision. Section III presents an exemplary use case of PERMIT. The overall PERMIT framework is portrayed in Section IV. Section V introduces the PERMIT Slice orchestration system and discusses its challenges. The paper concludes in Section VI.

II. NETWORK SLICING

The terminology “network slicing” has caught much attention in research communities and industries, as well as standards developing organizations (SDOs) such as 3GPP and ITU-T. Although the definition of “network slicing” is still under heavy discussion, it generally means an isolated collection of resources and functions implemented through software programs on top of the resources to enable QoS guarantee for the network requirements as well as in-network processing along end-to-end communications.

In PERMIT, network slicing (slicing here on) is considered to be one of the most important concepts to realize personalization of mobile network access of users. Although the slicing in mobile networking context has been often addressed for enabling different classes of communications, such as eMBB (enhanced Mobile Broad Band), mMTC (massive Machine Type Communications), URLLC

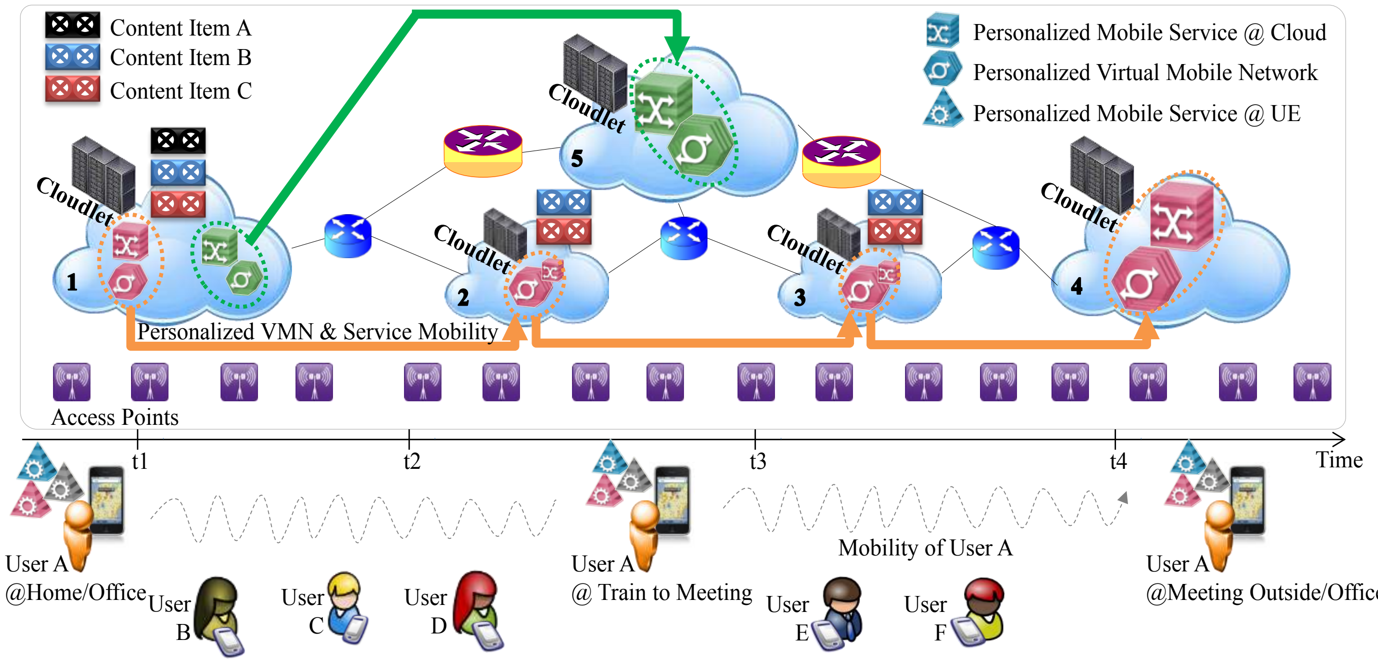

Fig. 1 — Personalized end-to-end mobile connection: use cases.

(Ultra Reliable and Low Latency Communications), in PERMIT, slices are instantiated per users, and in the most extreme case per device and/or per application.

It is a well-known fact that the concept of slice in networking has been first introduced in the overlay network research efforts, such as PlanetLab, in 2002. At that time, a slice has been defined as an isolated set of network bandwidth, computational, storage, resources allocated for a group of users that "program" network functions and services over their overlay network overlaid across "the planet". Since various network virtualization testbed efforts such as GENI, VNode, FLARE, Fed4Fire, have inherited the concept of slices as a basis of the infrastructures, as a set of programmable resources to tailor new network services and protocols, it is quite natural to use the slice concept in PERMIT to personalize the access and in-network edge processing for mobile network users.

III. PERMIT – Exemplary Use Case

Fig. 1 illustrates the basic concept behind PERMIT. A distributed architecture of micro- data centers, named cloudlets, being interconnected with a set of macro- data centers and forming a federated cloud, is assumed. Mobile users access their mobile services via a network of access points. Some mobile services can be accessed from cloudlets; others are accessed directly from their central servers. As conceptually considered by NFV [4], the mobile connectivity service of the users is managed through virtualized mobile core network slices, created as a composition of images running Virtualized Network Functions (VNFs), such as Mobility Management Entity and Packet Data Network Gateway on Virtual Machines (VMs) or containers in data centers forming Virtual Mobile Networks (VMN) [5][6].

In the PERMIT platform, prior knowledge on the trajectory of User A is assumed and that is through context awareness. For example, User A will be taking the train for a business meeting at a location different than his home, located next to Cloudlet1, passing by areas in the vicinity of Cloudlets 2 and 3. When being in the area of Cloudlet1, the envisioned PERMIT platform pushes services and content that User A is deemed to be interested in using on his way to the meeting or during the meeting. For the sake of energy efficiency, this operation is done when User A is charging his equipment. The "pushed" services and content are stored locally at User A's equipment so they can be accessed with no latency. While being at home, User A has his personalized VMN and some of his personalized mobile services hosted on VMs instantiated at Cloudlet1. As User A can also use the fixed network for part of his services, virtual resources running his VMN can be of small sizes in terms of allocated CPU, memory, storage, while using a large amount of network bandwidth. When User A leaves home and goes towards his meeting destination. The envisioned platform knows that based on User A's preferences, User A will be interested in watching content items A, B, and C that cannot be pushed and pre-stored at User A's User Equipment (UE) for example due to copyright, or because they are live event streams.

However, given the mobility features of User A, the platform recommends content items B and C as they are available at Cloudlets 2 and 3 that the user will be passing by. Furthermore, the platform may know that at the access points in the vicinity of Cloudlet 3, there may be not sufficient radio resources to service content item B with an acceptable QoE. The platform may then recommend content item C to User A, although it is the least preferred in comparison to content items A and B. In this manner, the mobility features of User A and the dynamics of the underlying communications infrastructure are reflected in the service personalization to positively impact the behavior of mobile users, directing them towards accepting mobile services that are, on one hand, deemed less congesting to the underlying infrastructure and on the other hand, that can be best serviced to the users given their mobility features and the current/future network conditions. While on the move, the personalized VMN of User A follows his mobility by having it migrate from Cloudlet 1 to Cloudlets 2, 3 and finally to Cloudlet 4 where the business meeting

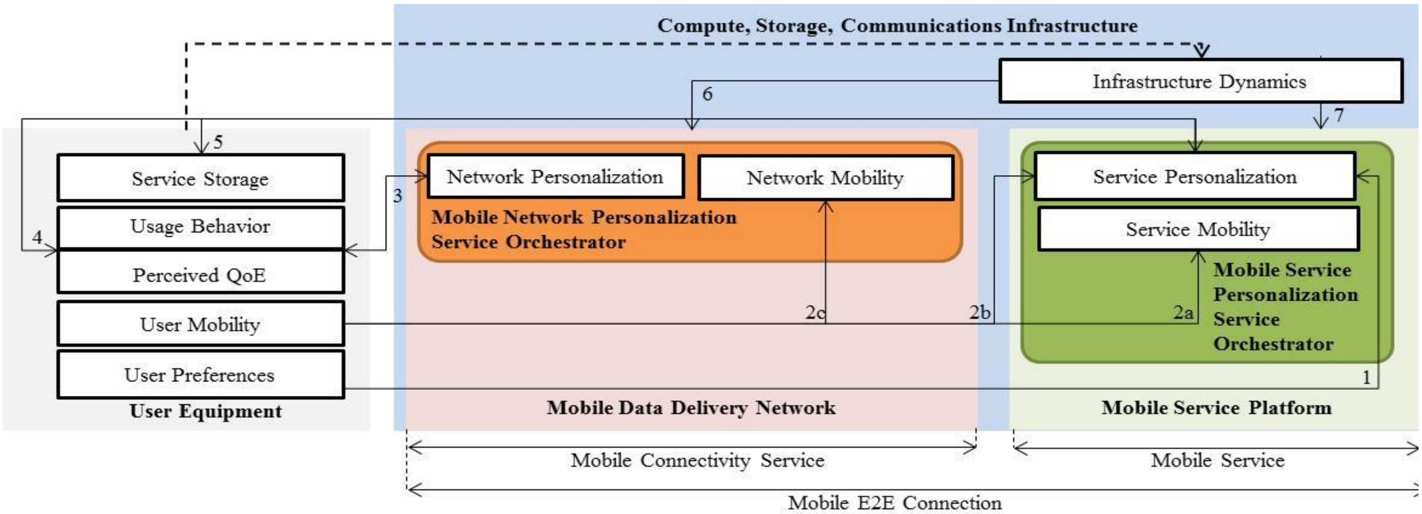

Fig. 2 – PERMIT VMN architecture.

will take place, and this is similar in spirit to the Follow Me Cloud concept [7].

As User A will need a larger mobile core network to process the services he intends using and showing to his business partner, User A's VMN slice gets instantiated on virtual resources with larger sizes. This is a good example that shows how a user may have his mobile network personalized to his needs and the requirements of his mobile services, ensuring that his mobile core network is always in his proximity.

Another PERMIT use case, depicted in Fig. 1, concerns Users B, C, and D. The three users belong to the same social network or to the same vertical business case such as cars of the same manufacturer. They frequently exchange among themselves messages and multimedia content. A VMN slice customized for the needs of Users B, C, and D and the requirements of their social network service is instantiated at Cloudlet 1, in the proximity of which the three users reside. As Users E and F become interested in joining the social network of Users B, C, and D, the respective VMN gets moved to Cloudlet 5 that is deemed optimal for the five users and is run on virtual resources with larger sizes to be able to process the increased social network traffic expected from the five peers. This is another good example to showcase how service-tailored mobile networking can be attained.

IV. PERMIT Framework

Fig. 2 schematically depicts the main components of the PERMIT framework. It mainly consists of two orchestrators, namely the Mobile Network Personalization Service Orchestrator (MNP-SO) and the Mobile Service Personalization Service Orchestrator (MSP-SO). PERMIT also envisions some changes to the user equipment as detailed below. The MNP-SO and MSP-SO entities can run separately or jointly on dedicated hardware or as software on VMs with adequate characteristics. These two entities incorporate all necessary intelligence for mobile service personalization and mobile network personalization, respectively. They are decision making entities that decide how mobile services and the lightweight VMN to transport them shall be personalized to the current and to the anticipated needs of a mobile user or a group of mobile users. In their decision making procedures, both MNP-SO and MSP-SO take into account the underlying infrastructure dynamics, processing and storages capabilities of user equipment, and users' contextual information like mobility and resource usage patterns.

In the envisioned PERMIT framework, lightweight VMNs are expected to run as a slice on one or multiple instantiated virtual resources [5][8]. These lightweight VMNs are expected to have the flexibility to serve one single service for one individual user, multiple services for one individual user, or a group of users. This flexibility can be attained by composing the lightweight VMN of uncorrelated building blocks that can be freely and dynamically combined or separated as per the requirements of the target mobile services and the needs of the serviced users [5].

Indeed, in PERMIT, a mobile network component (a network function), is defined in terms of its application logic and data, as depicted in Fig. 3. The application logic is decomposed into compute "blocks", including one basic block that provides minimal core network functionalities and of multiple added-value blocks, extending the basic components to provide additional network services such as communication reliability, Quality of Service and mobility support or accounting. The blocks run either at the network edge or in the cloud, depending on latency, bandwidth, resiliency and security requirements.

The composition of the blocks should follow the latest frameworks in service composition such as the ones based on micro-services APIs or based on event bus communication. Both of these ensure the loose coupling between the main functionality and the other modules, permitting the addition on demand, even during runtime of new functions, as well as a comprehensive separation of the liabilities. The PERMIT architecture will further optimize these mechanisms to address the end-to-end delay while processing a request through multiple compute blocks.

The PERMIT architecture enables the configuration of the in-service parameters related to the network functions as well as to the composition of the network functions within end-to-end services. This approach enables the creation of lightweight, customizable, and

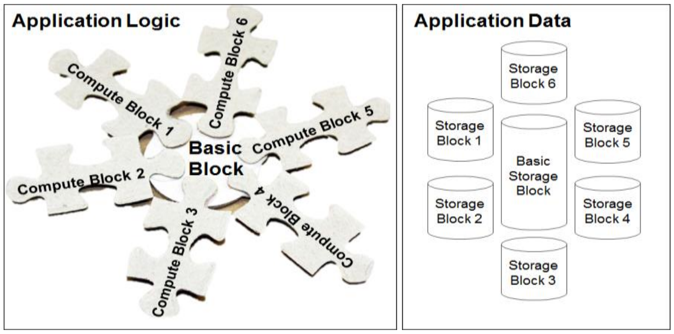

Fig. 3 – Network functions consisting of application logic and application data.

truly elastic mobile networks with network services/blocks that can be adapted to the users' needs. It will also enable seamless decomposition of application logic and data, allowing them to be moved to more convenient network locations. It is worth noting that careful attention shall be paid to how to gracefully compose the different functionalities within the network functions, especially considering the performance of multi-vendor software components, in order to provide a retraceable service level as well as to separate liabilities in case of failures.

At the protocol level, the lightweight feature of VMNs can be also achieved by simplifying a number of procedures typically required for mobile networks like Authentication, Authorization, and Accounting (AAA) and charging functions. Dependencies on data anchoring (i.e. Packet Data Network Gateway in Evolved Packet System) and mobility anchoring (i.e. Serving Gateway in EPS) concepts shall be relaxed if not completely replaced, as in the Follow Me Cloud concept [9]. With customizable VMN currently unsupported communication modes that mimic connectionless communication over shared media becomes possible over mobile networks.

In PERMIT, the personalization of mobile networks for a mobile user or a group of mobile users is achieved by anticipating the needs of the mobile services of this user or this group of users. Indeed, once the needs of a mobile service or a set of mobile services received by an individual user or a group of users are anticipated, the right VMN with the right characteristics (e.g., composing building blocks, total number of VMs involved, the locations of their respective DCs, their respective CPUs/memory/storage, etc.) can be identified so VMNs can scale up and down as per the assessed needs. Another aspect of VMN personalization consists in its mobility to a different data center when required. For this purpose, it is possible to leverage different algorithms and mechanisms [10] that decide on and enforce the VMN mobility as per the mobility patterns of the served mobile users and/or the dynamics of the underlying infrastructure, in a way that the "mobile network", serving a user or a group of users, follows their mobility. This decision may be based on several, possibly conflicting attributes/criteria such as the mobile service type (e.g., delay sensitive), the perceived Quality of Experience, the migration cost, the activity level of the users, the usage behavioral patterns, the mobility patterns, and the dynamics of the underlying communication infrastructure (e.g., for load balancing). Inputs, used for VMN personalization and VMN mobility, and relevant to users' mobile service usage behavior, perceived QoE, user's mobility, and dynamics of underlying communications infrastructure are schematically depicted in Fig. 2 through arrows 2c, 3 and 6.

The personalization of the mobile service depends first of all on the user preferences for the service delivery (arrow 1 in Fig. 2) and on its mobility (arrow 2). Based on insights on the user behavior and on its perceived QoE, the network is customized according to the user needs (arrow 3). As multiple users may have the same network requirements, the customization can be seen as a user classification problem towards the appropriate cluster of users which have the optimal handling of the communication requirements. Similar to the network customization, also a service customization may be executed (arrow 4). As the users may come from different customized networks, mainly due to their multi-application terminals, the customization of the applications should consider the customization of the network as a given parameter. A further step in the customization, is the distribution of the service data (arrow 5) to the UE when needed or when the network conditions are appropriate, depending on the specific user behavior.

Finally, the customization highly depends on the availability of the infrastructure resources for the specific customization (arrow 6 and arrow 7). The decisions of both the MNP-SO and MSO-SO depends on the possibility of the infrastructure to support at the specific location their needs. As the PERMIT architecture assumes a very large number of slices up to one for each network user, it is possible that the network infrastructure will not have enough momentary resources to handle the subscriber communication. In this case, the subscriber should be classified in a default communication class for which the processing is handled by a central data center within a default slice similar to the current network infrastructure.

V. PERMIT VMN SLICE ORCHESTRATION SYSTEM

PERMIT aims at achieving elasticity, flexibility, dynamicity, scalability, manageability, and efficiency beyond the current network level by building on-demand VMN slices, customized to the service requirements, through this highly reducing the network overhead.

PERMIT architecture is envisioned to function for different slices ranging from slices of individual users to IoT application slices as well as for verticals such as industrial control systems, autonomous driving, virtual reality or video streaming. The resulting system will consist of numerous network slices, running in parallel and being composed for end-to-end service delivery (Fig. 4).

Each slice consists of a set of virtual network functions within both the control and the data planes, customizable to the particular service types or vertical market needs or personalized to the individual end-user, as presented earlier. From the perspective of the orchestration, the VNF components of each slice are seen as software functions which may be composed and customized, thus transparent to the functions they handle.

The VNFs accommodate the intrinsic features of the slices and their changing requirements, such as scaling up to match sudden growth in their traffic or smooth mobility to another network location. Taking also into account the requirements of the service delivery in terms of latency, reliability and security together with the large number of slices, the resource control becomes highly complicated.

With the deployment of a large number of software services across multiple cloud and edge data centers, the complexity of the system increases beyond the capabilities of a single orchestration node. Moreover, a centralized network orchestrator may not be able to make the appropriate decisions and enforce them in due time, as it needs to handle a large amount of runtime operations, especially related to the sharing of the common data path environment. These delay and scalability limitations can be overcome through a distributed orchestration system where parts of the orchestration functionality are delegated to the edge nodes [12]. Data path sharing between the different slices and the decisions which require immediate response, such as network function failures, are particularly suitable for such a delegation.

Fig. 4 shows a high level architecture of the PERMIT VMN slice orchestration system. The physical infrastructure consists of hardware for computing, storage, networking and monitoring. Those equipment can be administrated by the same entity or could belong to different domains. The slice orchestration plane of the architecture include images of VNFs, which represent the software version of existing network equipment. These VNFs could consist of building blocks (Fig. 3) designed in a clean-slate fashion or as components of existing network equipment. The VNF slice orchestration system is the main component of the architecture. It creates slices of VNFs for an individual user or a group of end-users of a vertical. These slices can be created following pre-defined blueprints or in a fine-granular fashion, taking into account inputs relevant to end-users' mobile service usage behavior, perceived QoE, and users' mobility, as discussed earlier. The slice orchestration system can be owned by a cloud provider, a mobile operator, or a new stakeholder. The users of the system can be vertical providers (e.g., automotive and IoT service provider), a mobile application developer or an individual end-user wishing for personalized mobile telecommunication service. Users can communicate to the slice orchestration system via well-defined northbound interfaces (e.g., following OMA guidelines).

In order to separate the orchestration concerns, the following PERMIT orchestration levels are considered. First, a basic NFV resource orchestrator is considered, able to broker the available virtual resources, as provided by a Virtualized Infrastructure Manager (VIM) to the different slices. The NFV orchestrator receives resource allocation requests from the VNF managers, one for each slice which are aware of the specific slice logic. The role of the VNF manager is to transmit network function placement [10] and scaling requirements to the NFV orchestrator as well as to transmit to the VNFs the dependency parameters in order to enable the communication between the different VNF Components within a slice.

To reduce the complexity and to be able to manage appropriately the slice specific operations, a VMN slice orchestrator is added to the architecture having the role of orchestrating the functionality within the specific slice. This includes the acquisition of monitored data on the specific service agents within the slice, a composition logic for the VNF components, enabling the appropriate processing flow allocation according to the momentary available resources as well as the interaction to the other slices within the system.

For this, the management system as well as the communication plane are extended with a Slice Border Control (SBC), which similar to the Session Border Controllers in the current architecture, enables the filtering and the classification of both the inbound and outbound data traffic, application level firewall as well as the appropriate forwarding to the components within the slice or to the SBC of the peer slice.

Through this separation of concerns within multiple functions, the possible policy conflicts are mitigated, although due to the shared resources, it may happen that the exact requirements of a specific slice would be partially fulfilled by the infrastructure. Considering the large number of slices, the alternative to provide the information on available resources to each VNF manager is very complex.

From the users' perspectives, PERMIT will facilitate a fully personalized and elastic End-to-End (E2E) mobile connection service, which will provide mobile users with easy and efficient access to advanced mobile services. Indeed, with PERMIT, a mobile user may have his mobile network fully personalized to his current and anticipated needs or to the requirements of his mobile services. True service elasticity will be attained: services of heavy mobile phone users shall never be throttled. Fair charging models can be also achieved: light and medium mobile phone users will be having their respective VMNs running on smaller VMs which will enable them to be fairly charged for only what they have indeed consumed. A group of users of a vertical or receiving a particular service may have a mobile network fully customized to their needs and the requirements of their mobile service. This shall enable the much-desired service-tailored mobile networking concept. Furthermore, different VMNs with the right processing features can be created for different services as per the specifications of each service. Instead of being "locked-in" the same mobile network for the delivery of all service types, users will be then having the flexibility in subscribing to the most suitable VMN slice to receive a particular service type. Subscription to multiple VMNs for multiple service types accordingly becomes possible. This requires the UE to have the capability to simultaneously connect to and steer mobile traffic across multiple VMN slices, optimally created for a set of services (Fig. 5). A default connectivity slice will be always assumed for each user, being available for applications which use non-customized service delivery. A UE shall be able to discover existing VMN slices, and request the creation of a new one. The creation of a new VMN slice for a particular service can be UE-initiated (following a set of rules and policies) or network –controlled when discovering that the current slice used for the application is highly inefficient for the delivery of the service to the user or to a group of uses. A UE shall also have the ability to leave a VMN slice, join an existing one, or upgrade an ongoing one, either on a UE request or on a network based classification of the UE.

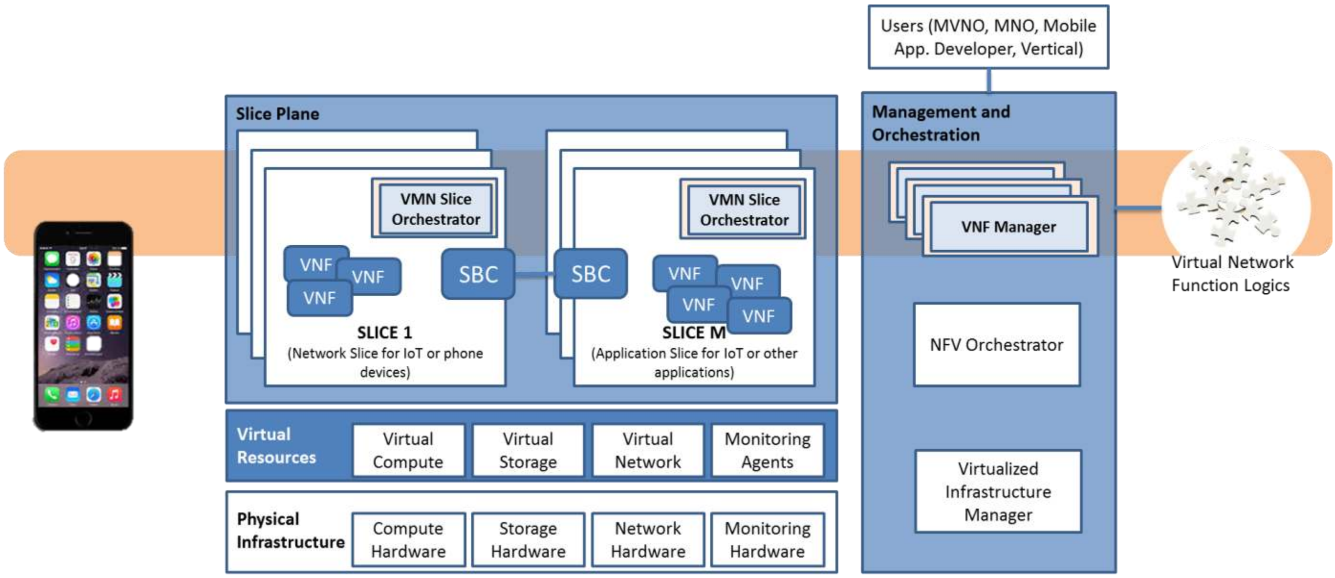

Fig. 4 – PERMIT VMN Slice Orchestration System: diverse VMN slices running in parallel and serving diverse verticals or individual users.

The slice concept can be further extended to support the slicing of resources within a single UE. In the current smartphone market, it is quite common that handhelds come with application vending facilities such as Google Play on android phones, iTunes App Store for iPhones, iPads, and iPods, and own application market places in the other smartphones. Applications on top of those handhelds must be examined and approved by handheld operating systems vendors, and must run on the sandboxes they prepared. We posit that the application market places and the sandboxes as execution environments should be implemented, being isolated within separate slices, so that one can have multiple, different, personalized execution environments and available application suites within the single hardware handheld. With this envisioned UE Slicing concept, one may have multiple personalized containers within a single handheld, so that one may have different security and privacy contexts such as private usage and corporate usage. Considering data contamination and privacy breaches often observed in the mixed use of a single handheld for private and public matters, it makes a lot of sense to introduce isolation between different usages. If multiple phones are carried to avoid such mishaps, it is reasonable to consolidate them into a single one using slicing techniques. Alternatively, one may have multiple different phone operating systems so that one may benefit from different application suites and execution environments. Embedded operating system virtualization and network virtualization technologies have already advanced to support such a concept of UE slicing [16][17]. However, to the best of the authors' knowledge, PERMIT is the first to consider end-to-end slicing, including UE slicing, mobile network slicing, edge computing and cloud computing.

In PERMIT, a user can have his personalized VMN and his personalized mobile services constantly following him. In this way, PERMIT will support a wide gamut of high quality services, customized to users' preferences, behavior, and mobility features. Emerging devices, such as Microsoft HoloLens, will largely benefit from the finding of PERMIT, particularly when their users are moving in dense smart cities on board smart connected vehicles.

PERMIT will also define novel and promising business opportunities for cloud providers, especially in the area of providing value-added services beyond the basic infrastructure sharing. It also represents an innovative and ambitious solution to open up mobile networks and revolutionize the mobile networking principles going form large scale ubiquitous uniform connectivity service to highly efficient, tailored to the specific needs, individual service support. In PERMIT, a new set of business stakeholders may also emerge, orchestrating the mobile service and mobile network personalization

Fig. 5 – VMN slicing from UE perspectives – A UE may have a default slice for basic connectivity service and additional slices, each optimally created for a mobile service.

for individuals and groups of users as well as orchestrating the interaction between the mobile service slices.

VI. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

In this paper, we propose the PERMIT approach that is expected to act as a catalyst for structural changes to the current communication system configuration, whereby both i) the mobile delivery network and ii) the mobile services it supports are personalized for each individual user, or alternatively customized for groups of users/verticals.

In PERMIT, service personalization goes beyond the traditional approaches whereby service personalization is based on the classical users' preferences, explicitly indicated by the users or deducted through collaborative filtering from other means – e.g., social networks resulting into a specific parametrization of the uniform network. In PERMIT, users' mobility patterns, their mobile service usage behavioral patterns, and the dynamics of the underlying communications infrastructure are all taken into consideration for acquiring both network and service personalization, treating service personalization and networking customization as two flexibility enabled and complementary components.

To always ensure short response times for emerging advanced mobile services, the mobility of i) mobile services, ii) the personalized mobile networks, or iii) both is enabled towards the proximity of the respective mobile users across the overall mobile service area, as per the mobility features of mobile users, and in a seamless and cost-efficient manner. This will make both the mobile delivery network and the mobile services – after being personalized – constantly following their respective mobile users.

With such an approach, service personalization and network personalization can take place dynamically and interactively. This also enables the much-desired service-tailored mobile networking concept, achieving a so far unprecedented level of flexibility in service-specific optimizations, fine-grained network resource slicing and transparent capacity scaling. In this regard, scalable programming of data plane and data paths and fine-grained mobility management of various services, considering both centralized and distributed approaches are needed. This defines a promising research area that will stimulate the relevant community of researchers.

ACKNOWLEDGEMENTS

This work was partially supported by the TAKE 5 project funded by the Finnish Funding Agency for Technology and Innovation (TEKES) and in part by the Finnish Ministry of Employment and the Economy. It is also partially supported by the European Union's Horizon 2020 research and innovation programme under the 5G!Pagoda project with grant agreement No. 723172.

REFERENCES

[1] Y. Jin et. al, "Characterizing Data Usage Patterns in a Large Cellular Network," in Proc. SIGCOMM CellNet Workshop, Helsinki, Finland, Aug. 2012.

[2] Y. Zhang and A. Arvidsson, "Understanding the Characteristics of Cellular Data Traffic," in Proc. SIGCOMM CellNet Workshop, Helsinki, Finland, Aug. 2012.

[3] T. Taleb, K. Samdanis, and A. Ksentini, "Supporting Highly Mobile Users in Cost-Effective Decentralized Mobile Operator Networks," in IEEE Trans. on Vehicular Technology, Vol. 63, No. 7, Sep. 2014, pp. 3381-3396.

[4] ETSI GS, "Network Function Virtualization (NFV) Management and Orchestration,
NFV-MAN 001 v. 0.8.1, Nov 2014.

[5] T. Taleb, M. Corici, C. Parada, A. Jamakovic, S. Ruffino, G. Karagiannis, and T. Magedanz, "EASE: EPC as a Service to Ease Mobile Core Network," in IEEE Network Magazine, Vol. 29, No. 2, Mar. 2015. pp.78 – 88.

[6] T. Taleb, "Towards Carrier Cloud: Potential, Challenges, & Solutions," in IEEE Wireless Communications Magazine, Vol. 21, No. 3, Jun. 2014. pp. 80-91.

[7] T. Taleb, A. Ksentini, and P. Frangoudis, "Follow-Me Cloud: When Cloud Services Follow Mobile Users", to appear in IEEE Transactions on Cloud Computing.

[8] T. Taleb, A. Ksentini, and A. Kobbane, "Lightweight Mobile Core Networks for Machine Type Communications," in IEEE Access Magazine, Vol 2, Oct. 2014. pp. 1128-1137

[9] A. Ksentini, T. Taleb, and F. Messaoudi, "A LISP-based Implementation of Follow Me Cloud," in IEEE Access Magazine, Vol 2, Oct. 2014. pp. 1340-1347

[10] T. Taleb, M. Bagaa, and A. Ksentini, "User Mobility-Aware Virtual Network Function Placement for Virtual 5G Network Infrastructure," in Proc. IEEE ICC 2015, London, UK, Jun. 2015

[11] A. Ksentini, M. Bagaa, T. Taleb, and I. Balasingham, "On using bargaining game for Optimal Placement of SDN controllers," in IEEE ICC'16, Kuala Lumpur, Malaysia, May 2016.

[12] A. Aissioui, A. Ksentini, A. Gueroui, and T. Taleb, "Elastic and Distributed SDN Controllers for Follow-Me Cloud," in IEEE Access Magazine, DOI 10.1109/ACCESS.2015.2489930, Vol. 3, Nov. 2015.

[13] P. Du and A. Nakao, "Application Specific Mobile Edge Computing through Network Softwarization," in IEEE CloudNet'16, Pisa, Italy, Oct. 2016.

[14] A. Nakao, P. Du and T. Iwai, "Application Specific Slicing for MVNO through Software-Defined Data Plane Enhancing SDN", IEICE Tran. on Communications, Vol. E98-B, No.11, pp.2111-2120, 2015.

[15] A. Nakao and P. Du, "Application and Device Specific Slicing for MVNO", The First Int'l Science and Technology Conf. on SDN & NFV, Moscow, Russia, Oct. 2014.

[16] C. Dall, et al, "The Design, Implementation, and Evaluation of Cells: A Virtual Smartphone Architecture", ACM Transactions on Computer Systems (TOCS), Volume 30, Issue 3, Aug. 2012

[17] J. Andrus et al, "Cells: A Virtual Mobile Smartphone Architecture", In. Proc. 23rd ACM Sym. on Operating Systems Principles (SOSP 2011), Oct. 2011