Towards 5G Network Slicing - Motivations and Challenges

By Prof. Alex Galis, *University College London*, and Dr. Chih-Lin I, *China Mobile Research Institute*

IEEE 5G Tech Focus: Volume 1, Number 1, March 2017

Abstract
This paper introduces the motivation for and the challenges of Network Slicing in the context of 5G Networks.

1. Introduction

5G networks are conceived as extremely flexible and highly programmable end-to-end connect-and-compute infrastructures that are both application- and service-aware, as well as being time-, location-, and context-aware.

These 5G networks represent:

- A revolution, over 4G networks, in terms of capacity, performance, and spectrum access in radio network segments; as well as
- A revolution of native flexibility and programmability conversion in all radio and non-radio 5G network segments: Radio Access Network, Fronthaul and Backhaul Networks, Access Networks, Aggregation Networks, Core Networks, Mobile Edge Networks, Software Networks, Software-Defined Cloud Networks, Satellite Networks and Edge IoT Networks.

5G networks enable new business opportunities by meeting the requirements of a large variety of use cases as well as enabling 5G to be future proof by means of: (i) implementing network slicing in a cost efficient way, (ii) addressing both end-user and operational services, (iii) supporting softwarization natively, (iv) integrating communication and computation, and (v) integrating heterogeneous technologies (including fixed and wireless technologies).

To take advantage of these opportunities, paradigm shifts and new mechanisms or updated mechanisms will be needed in all network domains. Equally important is the requirement for 5G networks to take a novel approach as to how to orchestrate, deploy, and manage services in 5G networks.

5G networks are expected to present a number of advantages. One in particular, is a high degree of flexibility. They enforce the necessary degree of flexibility, where and when needed, with regard to capability, capacity, security, elasticity, and adaptability. These networks will serve highly diverse types of communication – for example, between humans, machines, devices and sensors – with different performance attributes.

5G networks also represent a shift in networking paradigms: namely a transition from today's "network of entities" towards a "network of functions". Indeed, this "network of functions", and most likely, a "network of virtual functions", will in some cases result in the decomposition of current monolithic network entities into network functions. These functions will constitute the unit of networking for next generation systems, and should be able to be composed in an "on-demand", "on-the-fly" basis. A current research challenges consist of the architecture and design of solutions that identify the set of elementary functions or blocks from which to compose network functions, which are today implemented as monolithic elements.

Further advantages of 5G emerge in the areas of management, control of systems and resources. 5G networks enable uniform management and control operations that are becoming part of the design of dynamic software architectures. They can thereby host and execute services in one or more distinct network slices.

2. Network Slicing Context and Motivation

A number of definitions for network slicing as partitions of network resources were used in the last 10 years within the context of research into distributed and federated testbeds [1] and in future internet research [2]. More recently in 5G research revised definitions were used [3-7].

Network Slicing (NS) is an end-to-end concept covering all network segments. It enables the concurrent deployment of multiple logical, self-contained and independent shared or partitioned networks on a common infrastructure platform.

From a business point of view, a slice includes a combination of all the relevant network resources, functions, and assets required to fulfill a specific business case or service, including OSS, BSS and DevOps processes.

From the network infrastructure point of view, network slice instances require the partitioning and assignment of a set of resources that can be used in an isolated, disjunctive or non- disjunctive manner for that slice.

Currently Network Slicing refers to the managed partitions of physical and/or virtual network resources, network physical/virtual and service functions that can act as an independent instance of a connectivity network and/or as a network cloud. Network resources include connectivity, compute, and storage resources.

Network Slicing considerably transforms the networking perspective by abstracting, isolating, orchestrating, softwarizing, and separating logical network components from the underlying physical network resources and as such they enhance the network architecture principles and capabilities.

To support Network Slicing, the management plane creates a group of network resources (whereby network resources can be physical, virtual or a combination thereof), it connects with the physical and virtual network and service functions as appropriate, and it instantiates all of the network and service functions assigned to the slice. For slice operations, the control plane takes over governing of all the network resources, network functions, and service functions assigned to the slice. It (re-) configures them as appropriate and as per elasticity needs, in order to provide an end-to-end service. In particular, ingress routers are configured so that the appropriate traffic is bound to the relevant slice.

The establishment of slices is both business-driven as slices are the support for different types and service characteristics and business cases, and technology-driven as slices are a grouping of physical or virtual resources (network, compute, storage) which can act as a sub network and/or a cloud. A slice can accommodate service components and network functions (physical or virtual) in all of the network segments: access, core, and edge / enterprise networks.

Network operators can use NS to enable different services to receive different treatment and to allow the allocation and release of network resources according to the context and contention policy of the operators. Such an approach using NS would allow a significant reduction of the operations expenditure. In addition, NS makes possible softwarization, programmability and allows for the innovation necessary to enrich the offered services. Network softwarization techniques [5] may be used to realise and manage network slicing. NS provides the means by which the network operators can provide network programmable capabilities to both OTT providers and other market players without changing their physical infrastructure. NS enables the concurrent deployment of multiple logical, self-contained and independent, shared or partitioned networks on a common infrastructure. Slices may support dynamic multiple services, multi-tenancy, and the integration means for vertical market players (such as the automotive industry, energy industry, healthcare industry, media and entertainment industry, etc.)

3. Network Slicing Challenges for 5G Networks

In order to implement and use network slice functions and operations, there is a clear need to look at the complete life-cycle management characteristics of Network Slicing solutions based on the following architectural tenets:

- Underlay tenet: support for an IP-based underlay data plane the transport network uses to carry that underlay.
- Governance tenet: a logically centralized authority for all of the network slices in a domain.
- Separation tenet: slices may be independent of each other and have an appropriate degree of isolation from each other.
- Capability exposure tenet: allow each slice to present information regarding services provided by the slice (e.g., connectivity information, mobility, automaticity, etc.) to third parties, via dedicated interfaces and /or APIs, within the limits set by the operator.

In pursuit of solutions for the above tenets with the relevant characteristics within the context of 5G Networking, there is a need to address the following challenges and outcomes:

(1) A **Uniform Reference Model** for Network Slicing that describes all of the functional elements and instances of a network slice. It also describes shared non-sliced network parts.

(2) **Slice Templates**: Providing the design of slices to different scenarios. This outlines an appropriate slice template definition that may include capability exposure of managed partitions of network resources (i.e. connectivity compute and storage resources), physical and/or virtual network and service functions that can act as an independent connectivity network and/or as a network cloud.

(3) **Network Slice capabilities**, which are expected to be:

- Four-dimensional efficient slice creation with guarantees for isolation in each of the Data / Control / Management / Service planes. Having enablers for safe, secure and efficient multi-tenancy in slices.
- Methods to enable diverse requirements for NS, including guarantees for the end-to-end QoS of a service within a slice.
- Efficiency in slicing, specifying policies and methods to realize diverse requirements without re-engineering the infrastructure.
- Recursion, namely methods for NS segmentation allowing a slicing hierarchy with parent–child relationships.
- Customized security mechanisms per slice.
- Methods and policies to manage the trade-offs between flexibility and efficiency in slicing.
- Optimisation, namely methods for automatic selection of network resources for NS; global resource views; global energy views; Network Slice deployment based on global resource and energy efficiency; slice mapping algorithms.

- Monitoring the status and behaviour of NS in a single and/or muti-domain environment; monitoring of NS interconnection.
- Capability exposure for NS (allowing openness); with APIs for slice specification and interaction.
- Programmability and control of Network Slices.

(4) **Network slice operations**, which are expected to be:

- Slice lifecycle management including creation, activation / deactivation, protection, elasticity, extensibility, safety, sizing and scalability of the slicing model per network and per network cloud for slices in access, core and transport networks; for slices in data centres, and for slices in edge clouds.
- Autonomic slice management and operation, namely self-configuration, self-composition, self-monitoring, self-optimisation, self-elasticity for slices that will be supported as part of the slice protocols.
- Slice stitching / composition by having enablers and methods for efficient stitching / composition / decomposition of slices: vertically (through service + management + control planes); horizontally (between different domains as part of access, core, edge segments); or a combination of vertically + horizontally.
- End-to-end network segments and network clouds orchestration of slices [8-9].
- Service Mapping, by having dynamic and automatic mapping of services to network slices;

(5) **Efficient enablers and methods for integration** of the above capabilities and operations.

4. Status of Network Slicing in 5G Standards
3GPP SA2 has completed the first study item on the end to end architecture of 5G networks last month [6], in which network slicing is identified as the highest priority topic for which three key elements are defined:

Network Slice (NS): is a concept describing a system behaviour which is implemented via Network Slice Instance(s).

Network Slice Instance (NSI): is an instance created from a Network Slice Template (NST).

Network Slice Template (NST): is a logical representation of the Network Function(s) and corresponding resource requirements necessary to provide the required telecommunication services and network capabilities.

ETSI NFV issued NFV priorities on 5G Whitepaper [10] considering network slicing as the top feature to address. Various SDOs, industry fora, as well as open source community have their own definitions of network slicing. Their convergence from concept to practice will be key to fully realize 5G promise.

References

1. GENI Key Concepts - Global Environment for Network Innovations (GENI), http://groups.geni.net/geni/wiki/GENIConcepts.
2. Galis, A. et all - "Management and Service-aware Networking Architectures (MANA) for Future Internet" - Invited paper IEEE 2009 Fourth International Conference on Communications and Networking in China (ChinaCom09) 26-28 August 2009, Xi'an, China, http://doc.utwente.nl/73049/1/Management_and_Service-aware_Networking_Architectures.pdf.

3. Hedmar, P., Mschner, K., et all - NGMN Alliance document "Description of Network Slicing Concept", January 2016; ngmn.org/uploads/ media/160113_Network_Slicing_v1_0.pdf
4. Paul, M., Schallen, S., Betts, M., Hood, D., Shirazipor, M., Lopes, D., Kaippallimalit, J., - Open Network Fundation document "Applying SDN Architecture to 5G Slicing", April 2016; www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/Applying_SDN_Architecture_to_5G_Slicing_TR-526.pdf
5. Technical Report Application of network softwarization to IMT-2020, ITU-T FG IMT2020, December 2016, http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx
6. TR 23.799 V14.0.0, "Study on Architecture for Next Generation System" - Feb 2017; www.3gpp.org/ftp//Specs/archive/23_series/23.799/23799-e00.zip.
7. Galis, A., Makhijan, K., Yu, D. - Autonomic Slice Networking-Requirements and Reference Model, October 2016, https://tools.ietf.org/html/draft-galis-anima-autonomic-slice-networking-00
8. Guerzoni, R., Vaishnavi, I., Perez-Caparros, D., Galis, A., et al, "Analysis of End-to-End Multi Domain Management and Orchestration Frameworks for Software Defined Infrastructures , June 2016, onlinelibrary.wiley.com/10.1002/ett.3084/pdf
9. Karl, H., Dräxler, S., Peuster, M, Galis, A., et al - DevOps for Network Function Virtualization, July 2016, onlinelibrary.wiley.com/doi/10.1002/ett.3084/full
10. White Paper "Network Operator Perspectives on NFV Priorities for 5G", F 2017, http://portal.etsi.org/NFV/NFV_White_Paper_5G.pdf

Alex Galis (a.galis@ucl.ac.uk) is a Professor in Networked and Service Systems at University College London (UCL). He has co-authored 10 research books and more that 250 publications in the Future Internet areas: system management, networks and services, networking clouds, 5G virtualisation and programmability. He participated in a number of EU research projects including overall technical leadership of the MISA - Management of IP networks, FAIN – programmable networks, CONTEXT – context aware networking and AUTONOMIC INTERET – autonomic networking projects. He was a member of the Steering Group of the Future Internet Assembly (FIA) and he led the Management and Service–aware Networking Architecture (MANA) working group at FIA. He acted as PTC chair of 14 IEEE conferences including TPC co-chair of IEEE Network Softwarization 2015 (NetSoft 2015) and reviewer in more than 100 IEEE conferences. He is also a co-editor of the IEEE Communications Magazine feature topic on Advances In Networking Software. He acted as a Vice Chair of the ITU-T SG13 Group on Future Networking. He is involved in IETF and ITU-T SG13 network slicing activities and he is also involved in IEEE SDN initiative.

Chih-Lin I (icl@chinamobile.com) received her Ph.D. degree in electrical engineering from Stanford University. She has been working at multiple world-class companies and research institutes leading the R&D, including AT&T Bell Labs; Director of AT&T HQ, Director of ITRI Taiwan, and VPGD of ASTRI Hong Kong. She received the IEEE Trans. COM Stephen Rice Best Paper Award, is a winner of the CCCP National 1000 Talent Program, and has won the 2015 Industrial Innovation Award of IEEE Communication Society for Leadership and Innovation in Next-Generation Cellular Wireless Networks.

In 2011, she joined China Mobile as its Chief Scientist of wireless technologies, established the Green Communications Research Center, and launched the 5G Key Technologies R&D.

She is spearheading major initiatives including 5G, C-RAN, high energy efficiency system architectures, technologies and devices; and green energy. She was an Area Editor of IEEE/ACM Trans. NET, an elected Board Member of IEEE ComSoc, Chair of the ComSoc Meetings and Conferences Board, and Founding Chair of the IEEE WCNC Steering Committee.

She was a Professor at NCTU, an Adjunct Professor at NTU, and currently an Adjunct Professor at BUPT. She is the Chair of FuTURE 5G SIG, an Executive Board Member of GreenTouch, a Network Operator Council Founding Member of ETSI NFV, a Steering Board Member of WWRF, the ComSoc Rep of IEEE 5G Initiative, a member of IEEE ComSoc SDB, SPC, and CSCN-SC, and a Scientific Advisory Board Member of Singapore NRF. Her current research interests center around “Green, Soft, and Open”.

Editor: Geoffrey Ye Li