

On the Automation of RAN Slicing Provisioning and Cell Planning in NG-RAN

R. Ferrús, O. Sallent, J. Pérez-Romero, R. Agustí

Universitat Politècnica de Catalunya (UPC)

[ferrus, sallent, jorperez, ramon]@tsc.upc.edu

Abstract— Network slicing is a fundamental feature of 5G systems that facilitates the provision of particular system behaviours adapted to specific service/application domains on top of a common network infrastructure. While significant progress has already been achieved at specification level by 3GPP with regard to the functional support of network slicing, management solutions for the exploitation of these capabilities in the NG-RAN are still at a very incipient stage. In this context, this paper firstly presents a functional framework for the management of network slicing for a NG-RAN infrastructure, identifying the necessary information models and interfaces to support the dynamic provisioning of RAN slices. On this basis, the feasibility to automate the provisioning of RAN slices is discussed. Furthermore, a self-planning solution is presented to illustrate how a traditional network management process such as planning is expected to evolve to cope with the new challenges associated with RAN slicing management.

Keywords—Radio slicing; Network slicing; 5G New Radio; RAN slice; Network Slicing Management

I. INTRODUCTION

5G systems target the simultaneous support of a wide range of application scenarios and business models (e.g. automotive, utilities, smart cities, high-tech manufacturing) [1]. This expected versatility comes with a high variety of requirements on network functionalities (e.g. security, mobility, policy control features) and expected performance (e.g. peak rates above 10 Gbps, latencies below 1 ms with 10^{-5} reliability, 500 km/h mobility target) that cannot always be met through a common network setting. In this respect, the support for network slicing in 5G has become a foundational requirement to allow operators to compose and manage dedicated logical networks with specific functionality, without losing the economies of scale of a common infrastructure [2].

Each one of these logical networks is referred to as *network slice*, and can be tailored to provide a particular system behaviour (i.e. slice type) through the use of specific control plane and/or user plane functions to best support specific service/applications domains. For instance, a User Equipment (UE) for smart metering applications can be served through a network slice with radio access tailored to very small, infrequent messages and with no need to implement unnecessary functions (e.g. no mobility support). Similarly, a network slice can also be used to provide a particular tenant (i.e. an organization or business entity) with a given level of guaranteed network resources and isolation with regard to the operation of other concurrent slices. For instance, UEs/subscribers of a Public Safety agency can be served through a network slice that guarantees a minimum capacity during network congestion periods.

The realization of network slices considers, in the most general case, the support for specific features and resources both in the 5G Core (5GC) network part, referred to as Core Network slice, and in the New Generation Radio Access Network (NG-RAN) part, referred to as RAN slice. The realization of RAN slices is particularly challenging because it requires addressing how the pool of radio resources (i.e. RF bandwidth) available to one NG-RAN node (e.g. gNB) can be configured and operated to simultaneously deliver multiple and diverse RAN behaviours [3]. For example, a RAN slice could be purposely designed to provide specific RAN capabilities (e.g. Ultra Reliable and Low Latency Communications [URLLC] access) and network characteristics (e.g. coverage, capacity) according to the access part requirements of the Packet Data Unit (PDU) connectivity service associated with a particular network slice identifier (i.e. Single Network Slice Selection Assistance Information [S-NSSAI]) within a Public Land Mobile Network (PLMN).

The support of network slicing has been addressed in the 5G architectures of different 5G-PPP phase 1 research projects such as 5G-NORMA [4], METIS-II [5] or SESAME [6] and has become a central topic in most of on-going 5G-PPP Phase 2 projects. A complete solution for network slicing combines multiple facets, ranging from virtualization techniques for the abstraction and sharing of radio resources (e.g. network virtualization substrate concept in [7], descriptors for the configuration of RAN slices [3]) up to network slice lifecycle management solutions enabling the end-to-end automation across multiple resource domains (e.g. NESMO framework in [8]) and the delivery of *Network Slice as a Service* (e.g. 5G network slice broker concept in [9]). At specification level, 3GPP has already completed the first phase of the normative specifications regarding requirements and system architecture to support network slicing in 5G network [10][11]. However, the specification of a management framework for the exploitation of network slicing capabilities in 5G networks is still work in progress [12] and specific management solutions for the automated provisioning of network slices are at a very incipient stage, particularly for what concerns to the management aspects of network slicing in the NG-RAN.

In this context, and consistently with the holistic standardization effort conducted by 3GPP to support network slicing management in 5G systems, this paper provides three main contributions. First, a functional framework for the management of network slicing particularised for a NG-RAN infrastructure is presented, identifying the necessary information models and interfaces to support dynamic provisioning of RAN slices. Second, the feasibility to automate the provisioning of RAN slices is discussed in order to assess the degree of automation complexity. Third, a self-planning solution is worked out to illustrate how a traditional network management

process such as planning might evolve to cope with the new challenges associated with RAN slicing management.

The rest of the paper is organized as follows. Section II outlines the 3GPP work on the management of network slicing. Section III describes the network slicing management architecture for the NG-RAN. Section IV addresses the discussion on the provisioning automation and Section V elaborates on the self-planning applications for RAN slicing scenarios. Finally, concluding remarks are drawn in Section VI.

II. 3GPP WORK ON THE MANAGEMENT OF NETWORK SLICING

Support for network slicing in 5G systems has been already included in the first release of the normative specifications approved by 3GPP in December 2017 (Non-Standalone [NSA] 5G New Radio NR specifications). These specifications cover the functional aspects and protocols to deliver multiple network slices within a PLMN across the 5GC and NG-RAN functionality (e.g. S-NSSAI identifiers, standardised slice/service type [SST] values, procedures and functions for network slice selection, etc.).

Nevertheless, from a management perspective, the necessary specifications are still work in progress. According to TR 28.800, besides 5G network function management, 3GPP management system is recommended to support network slicing management for 5G. In this regard, 3GPP has already completed a study on management and orchestration of network slicing for next generation network (TR 28.801) and it is currently addressing normative specification work including: (1) network concepts, use cases and requirements for management of network slicing in mobile networks (TS 25.530); (2) management interfaces and procedures for provisioning of network slices (TS 28.531); (3) performance management and assurance data for network slicing (TS 28.550, TS 28.553, TS 28.554); (4) fault supervision for network slicing (TS 28.545) and (5) Network Resource Model (NRM) definitions for management of network slicing (TS 28.540, TS 28.542). In addition, a feasibility study is on-going to examine the business role models for network slicing in order to identify potential requirements that will enable a 3GPP system to adequately support those models.

A central aspect for the specification of the network slicing management system is the identification of the phases and high level tasks involved in the management aspects of a Network Slice Instance (NSI), which is defined as the set of network function instances and the required resources (e.g. compute, storage and networking resources) that form a deployed network slice. In this regard, Fig. 1 depicts the four phases and related tasks currently established in TS 28.530. The preparation phase includes all the tasks that are required to be done before a NSI can be created (e.g. network slice template design and on-boarding). The commissioning phase includes the creation task, in which the necessary resources are allocated and configured. The operation phase includes the activation, supervision, performance reporting (e.g. for KPI monitoring), modification, and de-activation of the NSI. When no longer needed, a NSI may be decommissioned (i.e. termination task) and resources released or reconfigured accordingly. An important principle laid out in TS 28.530 to build the network slice management functions (NSMF) set is the use of

standardized interfaces between the network slicing management functionality and the underlying network functions (or its network function related management functions) as well as between the different functions that could form part of the NSMF set. Of note is that while 3GPP TR28.530 establishes the general concepts, use cases and requirements for management of network slicing across the whole set of the network functions specified in 3GPP (e.g. NG-RAN, 5GC and IMS), the specifics of RAN slicing are not covered.

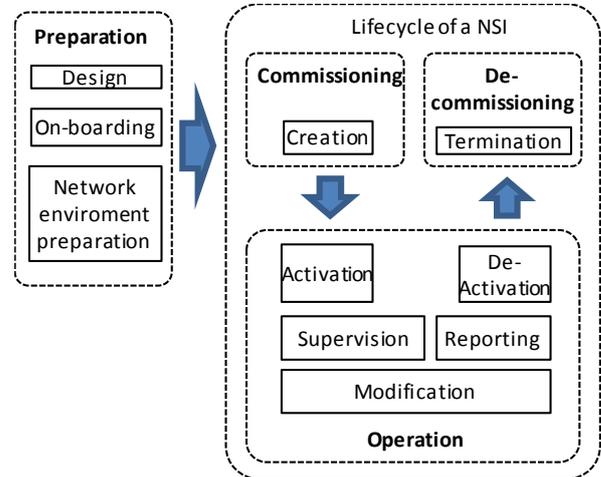


Fig. 1. Management aspects of a Network Slice Instance (NSI) (adapted from 3GPP TS 28.530)

III. FUNCTIONAL FRAMEWORK FOR MANAGEMENT OF NETWORK SLICING IN THE NG-RAN

This section presents a plausible functional framework for the realization of network slicing management within the NG-RAN. The functional framework, depicted in Fig. 2, is consistent with the general 3GPP management reference model for the Operations Support System (OSS) of a mobile network (3GPP TS 32.101) and well aligned with on-going standardisation work on management and orchestration of network slicing for next generation network discussed in the previous section. Details are given next, following a bottom-up approach.

A. Managed NG-RAN infrastructure

A NG-RAN infrastructure is to be formed by a collection of cell sites (i.e. NG-RAN Points of Presence [PoPs] where the antennas are located) and edge Data Centres (DCs), interconnected by means of a fibre and/or wireless-based Transport Network (TN). Focusing on 5G NR, the NG-RAN functionality for 5G NR access is delivered in the form of a single gNB NF that provides the full radio protocol stack functionality or split into a gNB-Centralized Unit (CU) NF and a gNB-Distributed Unit (DU) NF connected through a standardised interface (F1). These NFs can be implemented in dedicated hardware appliances (referred to as Physical Network Functions [PNF] in ETSI Network Function Virtualisation[NFV] terminology) or as Virtualized Network Functions (VNFs) running on a NFV Infrastructure (NFVI), as could be the most likely case for the gNB-CU. The placement of the gNB and gNB-CU NFs, virtualised or not, can be done at the cell sites or centralised in the edge DCs. In turn, gNB-DU

integration of such plethora of management applications within the NM layer is established by 3GPP, which only specifies the standardised interfaces and information models that are necessary for the interaction of the NM layer functionality with the rest of management components (e.g. Itf-N interface and NRM models for interaction with EM layer and ETSI interfaces and information models for interaction with ETSI NFV MANO).

On this basis, the support of RAN slicing management features within the NM layer enabling the dynamic creation of RSIs and leveraging NFV capabilities requires: (1) the specification of information models and interfaces to support dynamic provisioning of RAN slices and (2) the specification of information models to account for the manageable characteristics of RSIs inside the NM layer and to handle the interaction with EM and NFV management.

In order to accomplish (1), it is necessary the development of RAN slice templates to specify the service characteristics of a RAN slice and the support of provisioning interfaces through which RSI creation requests can be triggered by the Network Operator (NOP) or other network management systems. This entails the incorporation of a set of functions within the NM layer for the lifecycle management (LCM) of RAN slices. In this respect, the LCM functions, based on the service requirements expressed through RAN slice templates, will be in charge of orchestrating the necessary resources for the creation and operation of the RSIs. Importantly, the incorporation of this set of functions shall allow the NG-RAN NM layer to offer a RAN slice management service through a 3GPP standardised interface for: (i) Provisioning of RAN slices, enabling the creation of a new RSI or the use of an existing RSI, to satisfy specific requirements expressed through RAN slice templates and requested by the NOP or by an upper-level NS management function for the creation of an end-to-end NSI (e.g. RAN slice + 5GC slice); (ii) Performance management of a RSI, reporting performance measurement data of a RSI; and (iii) Report fault management of a RSI, reporting upwards fault management data of a RSI instance that requires NOP or upper-level system intervention.

In order to accomplish (2), new or extended NRM models that incorporate the necessary management objects for the configuration and operation of RSIs have to be defined. Also, the suitability of ETSI information models and interfaces shall be analysed to properly support the RAN slicing management needs within the NM layer when NFV is leveraged [15]. The interested reader is referred to [16] for more details on the use of NRM and NSD models for the management of RAN slices.

In addition to the incorporation of the RAN slicing LCM functions, the rest of management applications within the NM layer are also required to evolve to account for the new information models and management objects (e.g. performance management should be able to monitor KPIs per RSI for SLA compliance) as well as to properly interact with the LCM functions. In this regard, a compelling proposition is the development of a tight integration between the LCM functionality and NM functionality for cell planning and self-deployment, as further explored later on in this paper.

IV. COMPLEXITY OF AUTOMATION OF RAN SLICING PROVISIONING

The new 3GPP interface for RSI provisioning (shown in the management architecture depicted in Fig. 2) shall allow automating the provisioning processes of RAN slicing (i.e. from creation to termination of RSIs). The feasibility of such automation depends on the elements and configurations within the NG-RAN that could be potentially impacted for the realisation of the RSIs. As depicted in Fig. 3, if the deployment of a new RSI involves the installation of additional infrastructure or hardware modifications (e.g. add new cell sites to extend coverage or increase capacity or new Remote Radio Head [RRH] to support new RF bands), several actions with human intervention have to be necessarily conducted before the RAN slice could be provisioned in an automated manner. Otherwise, automation becomes feasible but with varying degree of complexity.

Automation is anticipated to be easily achievable with relatively low complexity if the provisioning of a RSI does not modify parameters that impact on the cell footprint but it is mainly based on the configuration of parameters such as network and network slice identifiers within cells (e.g. S-NSSAI(s), PLMN(s) that can be accessed in each cell), slice type selectors (e.g. selection of the radio access behaviour to be enforced, among a set of behaviours pre-configured/supported) and load guarantees and/or limitations to control the amount of traffic aggregate served by each RSI. A more complex case in terms of automation comes when provisioning also embraces the reconfiguration of cell parameters that modify to some extent the capacity/ coverage of the cells (e.g. transmission power, channel bandwidth, frequency band, mobility control configuration). In this case, automation feasibility mainly relies on the support of SON functionalities to re-adjust the modified cell footprint for optimised performance (e.g. re-computation of neighbouring cells, correction of coverage holes and cell overshoot situations, interference mitigation, mobility optimisation and so on). Ultimately, this reconfiguration of the cells might also involve the activation of new cells (e.g. new RF carriers), which could be served by the already operational gNBs (e.g. a gNB typically can handle multiple cells) or require the deployment of new VNF instances involving the NFV MANO systems (e.g. instantiation of a new gNB VNF to handle the new cell).

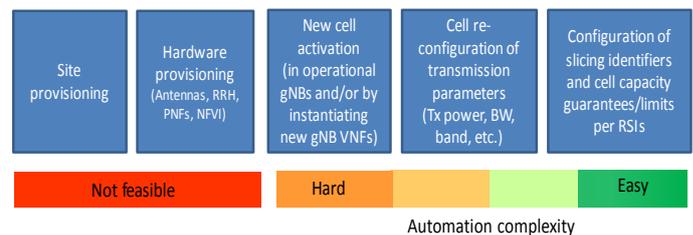


Fig. 3. Feasibility and complexity of automation in the RSI provisioning processes

V. TOWARDS SELF-PLANNING FOR RAN SLICING MANAGEMENT

A business environment where RSIs with diverse capacity, coverage and time scales can be dynamically provisioned as needed poses unprecedented challenges to the operator of the

NG-RAN in relation to traditional Radio Network Planning (RNP) in 2G/3G/4G. In turn, it is expected that 5G systems' deployment will stimulate multiple applications in numerous vertical markets and, therefore, the RAN will be required to respond to high diversified radio access needs in space and time. To ensure speedy and resource-efficient deployment of services, traditional RNP has to be evolved toward new models where specific coverage and traffic characteristics of RAN slices, both forecasted and in operation, shall be considered [17][18][19].

The creation of a new RSI may require some preparatory planning and resource deployment actions before its provisioning in order to meet specific space/time traffic and performance characteristics associated with the new slice. For example, the business feasibility of a RSI requested by a service provider in the automotive sector for in-car HD video services might be conditioned to the need to achieve higher peak bit rates than currently supported over main roads so that the NOP might have to necessarily upgrade the network accordingly before the RSI can be provisioned.

Nevertheless, from a capacity dimensioning perspective, it is anticipated that, given that in many cases the capacity required by a new RAN slice will grow progressively over time because the penetration in the number of customers takes time, a likely practice for the NOP is to rely on a more reactive approach (i.e., network capacity will be progressively increased at the pace demanded by the actual traffic). In this way, new application/services could be enabled from day zero even if the deployed capacity is not enough to meet long-term traffic expectations, and the NOP may rely on smart planning tools to optimise the upgrade process in which the NOP detailed knowledge on the actual use of network resources by the new services can be leveraged.

In this respect, Fig. 4 presents a functional framework for self-planning of the radio access capacity of a NG-RAN. Under this framework, the amount of deployed capacity to comply with traffic load and KPI requirements of the provisioned RSIs is checked in order to trigger particular planning actions such as changing cell transmission parameters (e.g., channel bandwidth modification, adding/removing carriers). Its operation involves two processes: capacity compliance and capacity dimensioning. The former considers a simplified model of the underlying network capacity while the later considers more detailed analysis through the interaction with more powerful radio planning and optimisation tools. Whenever the capacity compliance process envisages capacity shortages, the capacity dimensioning process is triggered. Depending on the degree of network automation and the policies indicated by the NOP for the self-planning application, the outputs of the capacity dimensioning process could directly trigger automatic cell re-configuration mechanisms or be escalated up to the NOP (e.g. the NOP is notified about network infrastructure shortages that could put at risk capacity compliance).

As illustrated in Fig. 4, the self-planning application interacts with the network inventory application within the NM layer to exploit live network configuration data such as detailed configurations of sites and cells and the actual configuration of the provisioned RSIs. Likewise, interactions with the PM/monitoring applications allow the self-planning application

to have access to live network measurements including KPIs derived from performance counters and geo-located mobile-traces. The measurement data is expected to be available with different granularity and scope (e.g. cell level and geo-located measurements, global and per-RSI measurements). The traffic load estimation module provides the input traffic to be considered in the radio planning and optimisation tool and in the capacity compliance process. It can be based on actual traffic extracted from the monitored live data or can include a traffic forecasting technique, in order to enable a predictive dimension to the self-planning process.

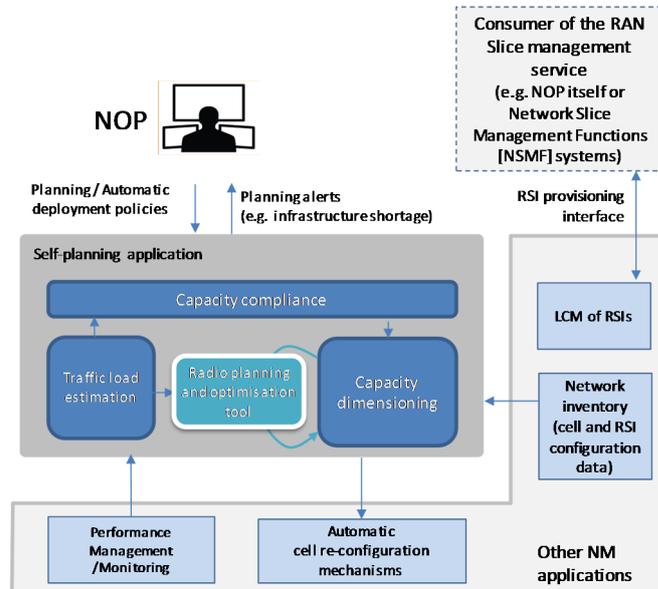


Fig. 4. Self-planning for RAN slicing management

VI. CONCLUDING REMARKS

Consistently with the holistic standardization effort conducted by 3GPP to support network slicing management in 5G systems, this paper has elaborated on a plausible functional framework for the realization of network slicing management in the NG-RAN, discussing the feasibility to automate the provisioning of RAN slices and highlighting the need to evolve the network planning solutions.

It has been anticipated that the automation of RAN slicing provisioning, which is realized at NM level, is achievable with relatively low complexity if the provisioning process involves the configuration of cell parameters such as network slice identifiers, slice type selectors and thresholds for load guarantees/limitations to control the amount of traffic aggregate per RSI. Otherwise, automation feasibility is highly dependent on the support of SON functionalities able to properly handle potential cell footprint modifications that might be triggered from the provisioning processes, including the automatic deployment of new cells.

Furthermore, the challenges that the dynamic provisioning of RAN slices poses to other network management applications that form part of the NM layer such as network planning, have been tackled in this paper by proposing a new functional framework for self-planning of the radio access capacity of a NG-RAN. The solution exploits detailed knowledge on the actual and/or forecasted use of network resources by the

provisioned RAN slices in order to drive automatic cell re-configuration mechanisms and/or provide notifications about network infrastructure shortages.

ACKNOWLEDGEMENT

This work has been supported by the EU funded H2020 5G-PPP project 5G ESSENCE under the grant agreement 761592 and by the Spanish Research Council and FEDER funds under SONAR 5G grant (ref. TEC2017-82651-R).

REFERENCES

- [1] NGMN Alliance, "5G White Paper", February 2015.
- [2] 3GPP TR 22.864: "Feasibility Study on New Services and Markets Technology Enablers - Network Operation; Stage 1 (Release 15)", September 2016.
- [3] R. Ferrús, O. Sallent, J. Pérez-Romero, R. Agustí, "On 5G Radio Access Network Slicing: Radio Interface Protocol Features and Configuration", IEEE Communications Magazine, Early access in IEEEXplore, January 2018.
- [4] P. Rost et al., "Mobile network architecture evolution toward 5G," in IEEE Communications Magazine, vol. 54, no. 5, pp. 84-91, May 2016.
- [5] METIS II White Paper "Preliminary Views and Initial Considerations on 5G RAN Architecture and Functional Design", March 2016.
- [6] O. Sallent, J. Perez-Romero, R. Ferrús, R. Agusti, "On Radio Access Network Slicing from a Radio Resource Management Perspective", IEEE Wireless Communications, October, 2017, pp. 166-174.
- [7] X. Costa-Perez, J. Swetina, T. Guo, R. Mahindra and S. Rangarajan, "Radio access network virtualization for future mobile carrier networks," in IEEE Communications Magazine, vol. 51, no. 7, pp. 27-35, July 2013
- [8] A. Devlic, A. Hamidian, D. Liang, M. Eriksson, A. Consoli and J. Lundstedt, "NESMO: Network slicing management and orchestration framework," 2017 IEEE International Conference on Communications Workshops (ICC Workshops), Paris, 2017, pp. 1202-1208
- [9] K. Samdanis, X. Costa-Perez and V. Sciancalepore, "From network sharing to multi-tenancy: The 5G network slice broker," in IEEE Communications Magazine, vol. 54, no. 7, pp. 32-39, July 2016.
- [10] 3GPP TS 22.261 v15.0.0, "Service requirements for the 5G system; Stage 1 (Release 15)", March 2017.
- [11] 3GPP TS 23.501 v15.0.0, "System Architecture for the 5G System; Stage 2 (Release 15)", December 2017.
- [12] 3GPP TS 28.530 v0.5.0 "Management of network slicing in mobile networks; Concepts, use cases and requirements (Release 15)", February 2018.
- [13] 3GPP TS 32.102 v14.0.0: "Telecommunication management; Architecture", March 2017.
- [14] 3GPP TR 32.864 v15.0.0, "Study on management aspects of virtualized network functions that are part of the New Radio (NR) (Release 15)", January 2018.
- [15] ETSI GS NFV-EVE 012 (V3.1.1): "Network Functions Virtualisation (NFV) Release 3; Evolution and Ecosystem; Report on Network Slicing Support with ETSI NFV Architecture Framework", December 2017.
- [16] R. Ferrús, O. Sallent, J. Pérez-Romero, R. Agustí, "Management of Network Slicing in 5G Radio Access Networks: Functional Framework and Information Models", submitted to IEEE Communications Magazine. Pre-print version available on arXiv.org, February 2018.
- [17] P. Muñoz, O. Sallent, J. Pérez-Romero "Capacity Self-Planning in Small Cell Multi-Tenant 5G Networks", IFIP/IEEE International Symposium on Integrated Network Management - 2nd International Workshop on Management of 5G Networks, Lisbon, Portugal, May, 2017.
- [18] P. Muñoz Luengo, O. Sallent and J. Perez-Romero, "Self-Dimensioning and Planning of Small Cell Capacity in Multi-Tenant 5G Networks," in IEEE Transactions on Vehicular Technology, vol. PP, no. 99, pp. 1-1, January 2018.
- [19] F. Bahlke, O. D. Ramos-Cantor, S. Henneberger, M. Pesavento, "Optimized Cell Planning for Network Slicing in Heterogeneous Wireless Communication Networks", available at <https://arxiv.org/pdf/1801.07067.pdf>, January, 2018.