

Marco funcional para la gestión de *slicing* en la red de acceso radio 5G

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

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

Departamento de Teoría de la Señal y Comunicaciones, Universidad Politécnica de Cataluña

Abstract- El soporte de *slicing* en los futuros sistemas 5G persigue acomodar, en una misma infraestructura de red, una diversidad de aplicaciones y modelos de negocio mucho mayor que el permitido actualmente con los sistemas 4G. Si bien el soporte de *slicing* a nivel funcional y de procedimientos ya ha sido cubierto en la primera entrega de especificaciones 5G, el desarrollo de soluciones para la gestión de las capacidades de *slicing*, particularmente en lo que concierne a la red de acceso 5G, se encuentra en una fase mucho más incipiente. En este contexto, esta comunicación presenta un marco funcional para la gestión de *slicing* en la red de acceso radio y proporciona un conjunto de ejemplos de la operativa de dicho marco funcional en un escenario ilustrativo.

I. INTRODUCCIÓN

A diferencia de los sistemas actuales 4G, diseñados para proveer una solución del tipo '*one size fits all*' para los servicios de conectividad móvil de banda ancha, se espera que los sistemas 5G sean capaces de soportar simultáneamente un amplio rango de escenarios de aplicación y modelos de negocio, resultantes de la que se anticipa como adopción pervasiva de las tecnologías 5G en diferentes segmentos de mercado (p.ej., automoción, *e-health*, *utilities*, *smart cities*, industria, agricultura, audiovisual y entretenimiento) y la introducción de modelos de provisión de servicio ágiles y flexibles (p.ej., Network as a Service, customización a través de la exposición de capacidades de red) [1]. Esta esperada versatilidad se presenta juntamente con una elevada variedad de requerimientos en las funcionalidades de la red (en términos p.ej., de seguridad, movilidad, tarificación, *policy control*) y en las prestaciones esperadas (p.ej., velocidades de transmisión de pico de 20Gbps / 10Gbps para el enlace de bajada / subida, latencias inferiores a 1 ms con niveles de fiabilidad en tasa de error de bit de 10^{-5}) que no siempre se podrán conseguir al mismo tiempo y con una misma configuración de la red.

Esta diversidad de requisitos sobre las funcionalidades de la red, prestaciones y modelos de negocio se sustentarán sobre el denominado *network slicing*, concepto que emerge como uno de los pilares sobre los que construir el ecosistema 5G. Este concepto permite proporcionar diferentes comportamientos de red en términos de funciones y prestaciones, referidos como *network slices*, sobre una misma red 5G. El diseño de una *network slice* puede obedecer al soporte de una aplicación específica que demande un comportamiento concreto de la red así como a la necesidad de proporcionar servicio con determinadas garantías a los terminales/usuarios de una organización o entidad concreta (referida como *tenant*).

Los aspectos de arquitectura funcional (i.e. entidades lógicas, identificadores, procedimientos) para el soporte de

network slicing en 5G ya han sido establecidos en la primera versión de especificaciones aprobadas por 3GPP en diciembre 2017 [2][3]. Asimismo, la implementación de *slicing* en redes de acceso radio ha sido objeto de estudio desde múltiples ángulos, incluyendo desde aspectos de gestión de recursos radio (RRM) hasta la explotación de técnicas de virtualización [4]-[8]. Sin embargo, la concepción y desarrollo de soluciones de gestión de *slicing* que permitan un aprovisionamiento dinámico y automático de *slices* [9] se encuentra en un estado mucho más incipiente, especialmente en lo que concierne a la red de acceso radio 5G, denominada NG-RAN (Next Generation - Radio Access Network) [10][11].

En este contexto, el objetivo de esta comunicación es presentar un marco funcional para la gestión de *slicing* en NG-RAN, detallando las funciones, modelos de información e interfaces necesarias para una gestión automatizada de la provisión de *RAN slices*. Estos aspectos se incluyen en la sección II. Posteriormente, a partir de una configuración de referencia ilustrativa de lo que podría ser un despliegue de una infraestructura NG-RAN, en la sección III se proporcionan ejemplos que describen la operativa de funcionamiento del marco funcional expuesto.

II. MODELO FUNCIONAL PARA LA GESTIÓN DE *NETWORK SLICING* EN NG-RAN

El marco funcional propuesto en esta sección e ilustrado en la Fig. 1 es consistente con los planteamientos del trabajo actualmente en curso en 3GPP en cuanto a la especificación de las capacidades de gestión y orquestación de *network slices* extremo a extremo en redes 5G [12]. La propuesta se desarrolla centrándonos primero en los elementos de la infraestructura NG-RAN cuya gestión se ve impactada por el soporte del *slicing*, y abordando posteriormente los componentes del marco de gestión.

A. Infraestructura de NG-RAN

Una infraestructura NG-RAN abarca un conjunto, normalmente elevado, de puntos de presencia (PoP)s donde se ubican las antenas y sistemas de RF (p.ej. cabezales radio remotos RRH), los equipos radio, los equipos de red de transporte (TN) y, de una forma cada vez más creciente, plataformas de computación para soportar capacidades de *edge computing*. Centrándonos en el acceso 5G New Radio (NR), la funcionalidad de estación base se denomina gNB, definido como una función de red (NF) con un conjunto de interfaces estandarizadas que alberga toda la funcionalidad necesaria para interactuar con los equipos de usuario (UE) a través de la nueva interfaz 5G NR. De cara a facilitar diferentes tipos de despliegue de la funcionalidad gNB, 3GPP ha definido una interfaz interna que permite desagregar el gNB

en una NF centralizada denominada gNB-CU, que alberga la parte superior de la torre de protocolos de la interfaz radio, y una NF distribuida denominada gNB-DU, a cargo de los protocolos de las capas inferiores. Estas diferentes NFs (gNB, gNB-CU, gNB-DU) pueden implementarse mediante hardware dedicado (denominándose en este caso NF físicas [PNFs]) o como NF virtualizadas (VNF) que se proporcionan como paquetes de software (p.ej., imágenes de máquinas virtuales) y pueden ejecutarse en plataformas de computación con hardware de propósito general (i.e., Network Function Virtualisation Infrastructure [NVFI]), como podría ser el caso más probable de implementación de una gNB-CU. La ubicación de dichas NFs, virtualizadas o no, puede realizarse de forma distribuida en los mismos emplazamientos donde están los sistemas de RF y antenas o bien llevarse a cabo de forma centralizada en determinados PoPs especialmente preparados para el soporte de plataformas de computación más potentes. Desde una perspectiva de configuración, la implementación de un RAN Slice, denominado en lo subsiguiente como *RAN Slice Instance* (RSI), puede admitir diferentes posibilidades en cuanto a cómo se organizan las funciones y los recursos de la infraestructura NG-RAN para su provisión, incluido cómo se distribuye o comparte el espectro de radio entre RSIs [7][8].

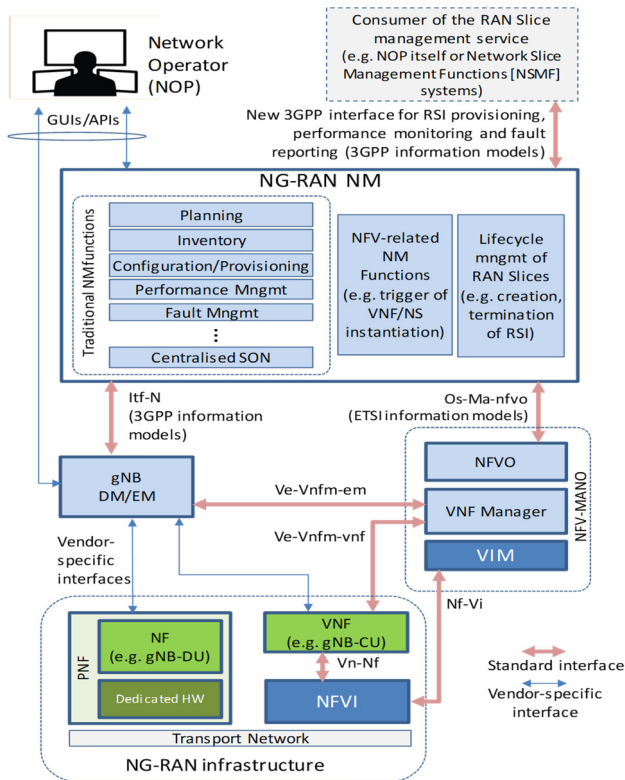


Fig. 1. Marco funcional para la gestión de *slicing* en NG-RAN

B. Gestión de los recursos de virtualización NFV

La gestión del provisionado y ejecución de las funciones de red gNB / gNB-CU / gNB-DU implementadas como VNF se realizaría a través de soluciones basadas en el marco ETSI Network Function Virtualisation (NFV) Management and Orchestration (MANO). Dicho marco está compuesto fundamentalmente por tres elementos: el orquestador de NFV (NFVO), el gestor de las VNF (VNFM) y el gestor de la infraestructura virtualizada (VIM).

La integración de la solución ETSI NFV MANO con el resto de componentes de gestión se sustenta en el despliegue de las interfaces definidas por ETSI tales como Os-Ma-Nfvo, Ve-Vnfm-em y Ve-Vnfm-vnf, tal y como se refleja en la Fig. 1. De esta forma, la creación, modificación o eliminación de servicios de red (NS) y VNF se puede desencadenar desde las capas NM / EM (explicadas a continuación) utilizando los modelos de información y procedimientos estandarizados en ETSI MANO.

C. Gestión de los elementos de red (EM)

La gestión FCAPS (Fault, Configuration, Accounting, Performance, Security) a nivel de NF y adaptada de forma específica de la funcionalidad proporcionada por los gNB se lleva a cabo a través de sistemas y herramientas de la denominada capa EM (Element Management). Los sistemas EM se basan principalmente en el uso de interfaces propietarias para interaccionar con las NF administradas dado que el fabricante que proporciona la NF es también habitualmente quien proporciona su sistema EM. Por otro lado, 3GPP especifica un conjunto de interfaces estandarizadas, denominadas colectivamente como Itf-N [13], para llevar a cabo la integración de la capa EM con la capa de NM (Network Management).

Dado que la capa de NM está concebida sobre principios de neutralidad en cuanto al proveedor de la capa EM, la interfaz Itf-N se basa en modelos de objetos que representan las características funcionales y recursos de los diferentes elementos de red desde una perspectiva de gestión. Dichos modelos se denominan Network Resource Models (NRMs).

D. Gestión de red (NM)

La capa de NM para el acceso radio puede incluir diversas aplicaciones de gestión, tales como planificación de red (p.ej. modelado y dimensionado de celdas), aprovisionamiento (p.ej. configuración de los parámetros operativos), funciones SON (Self Organising Networks) centralizadas (p.ej. ajuste automático de parámetros para la optimización de la movilidad), gestión de fallos de red (p.ej. correlación de alarmas entre elementos de red), gestión del rendimiento de red (p.ej. monitorización de red), etc.

El soporte de *slicing* requiere actualizar estas funciones para incorporar el concepto de RSI en el ámbito que les compete (p.ej., las herramientas de planificación de red deben ser capaces de hacer estudios por *slice*; la monitorización de KPI pasa a ser por *slice*, etc.). Además, resultante de la introducción de tecnologías de virtualización NFV en la NG-RAN y de la posibilidad de provisionar RSIs en base al despliegue dinámico de VNFs, la capa NM tiene que incorporar también funciones que posibiliten la instalación dinámica y flexible de la funcionalidad NG-RAN implementada por medio de VNFs, así como el control de su operativa. Si bien en la última versión completada de las especificaciones 3GPP (Release-14) se incluye ya un marco funcional de gestión para el soporte de virtualización en la red troncal y en la plataforma de servicios IMS, los aspectos específicos del soporte de virtualización en la RAN todavía son objeto de estudio [14].

La gestión de *slicing* en la RAN también requiere incorporar nuevas funciones en la capa NM para poder realizar la gestión del ciclo de vida (LCM, por sus siglas en inglés) de las RSI (p.ej. creación, modificación, terminación de RSIs).

En este sentido, uno de los elementos clave es el soporte de modelos de información e interfaces de aprovisionamiento que permitan al gestor de la infraestructura RAN (denominado como NOP en la Fig. 1), o bien a otro sistema de gestión de *slices* extremo a extremo que tenga que interaccionar con el NM para la gestión del segmento de *slice* en la NG-RAN, proceder a la creación/modificación/terminación de una RSI a través de la especificación de unos descriptores (denominados *RAN Slice Templates*) que contengan las características y la configuración de recursos asignados a cada RSI.

La automatización del proceso de aprovisionamiento es fundamental en escenarios donde la creación de *slices* pueda ser bastante frecuente. Además de la capacidad de aprovisionamiento automático, la interfaz proporcionada por la capa NM podría utilizarse también para exponer información de la fase operativa de una RSI (p.ej. monitorización del cumplimiento de los acuerdos de nivel de servicio [SLA], gestión y escalado de fallos en una RSI).

Finalmente cabe destacar que el soporte de *slicing* también requiere de la introducción de modelos de NRM nuevos y / o ampliados para la interacción de la capa NM y EM que incorporen los objetos de gestión necesarios para la configuración y el funcionamiento de los RSI, tal y cómo se elabora más en detalle en [15].

III. EJEMPLOS DE APLICACIÓN

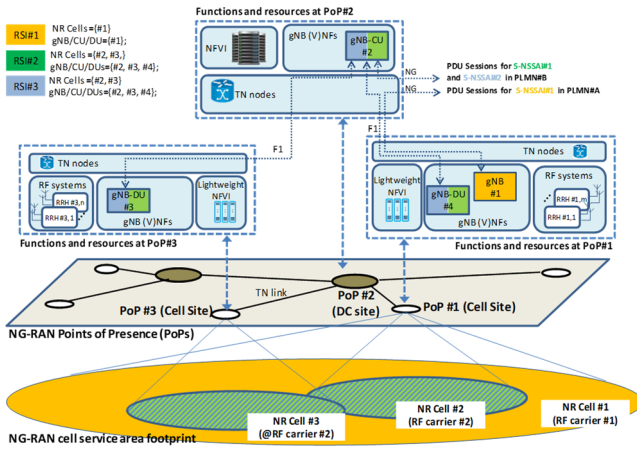
A partir de una configuración de referencia ilustrativa de lo que podría ser un despliegue de una infraestructura RAN con soporte de *RAN slicing*, esta sección describe la operativa de funcionamiento del marco funcional expuesto en la sección anterior bajo dos ejemplos de aplicación.

La configuración de referencia se muestra en la Figura 2a. En ella pueden verse diferentes PoPs de la infraestructura RAN donde se alojan los equipos y funciones descritas en el apartado anterior (i.e. equipos RF y hardware específico para soportar parte de la funcionalidad radio de acceso, infraestructura NFV, funciones gNB/gNB-CU/gNB-DU proporcionadas como VNF, equipos de comunicación para la interconexión de transporte entre RAN PoPs). La infraestructura RAN de referencia ilustra un escenario de operador neutro (i.e. gestor de la infraestructura RAN) que ofrece servicios de *RAN slicing* a diferentes proveedores de servicio externos que cuentan con redes troncales 5G (5G Core, 5GC) propias tales como operadores móviles tradicionales (MNO, por sus siglas en inglés), operadores de redes privadas y proveedores de servicios (SP) especializados en diferentes sectores. Por un lado, se considera que RSI#1 es una *RAN slice* configurada para servicios mMTC (p.ej. aplicaciones de telemetría) ofrecidos por un SP en el sector IoT mediante una red privada 5GC identificada como PLMN#A y un identificador de *network slice* (i.e. Single Network Slice Selection Assistance Information [S-NSSAI]) dado por S-NSSAI#1. Por otro lado, se considera que RSI#2 y RSI#3 ofrecen ambas acceso a la misma red pública 5GC de un MNO identificada como PLMN#B, si bien RSI#2 se utiliza asociada con el identificador S-NSSAI#1 para la provisión de servicios de banda ancha móvil (eMBB) a particulares y empresas, mientras que RSI#2 se asocia a S-NSSAI#2 y ofrece servicios de misión crítica (MC) para agencias en el sector de la seguridad pública y emergencias (PPDR, por sus siglas en inglés). En cuanto a los detalles de la configuración radio, se

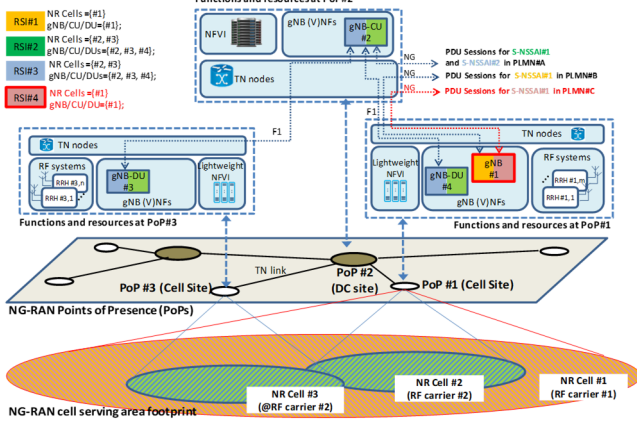
considera que RSI# tiene asignado un canal frecuencial dedicado a esta *slice* (RF Carrier#1 en la Figura 2a que podría ser, p. ej., un canal con el ancho de banda mínimo de 5 MHz soportado en NR en la banda de 700 MHz) que se explota en una macro celda (NR Cell#1) proporcionada mediante una gNB#1 VNF instanciada en la infraestructura NFV del PoP#1. Dicha configuración radio de RSI#1 se asume que es apropiada para los requisitos del IoT SP, que pretende conectar múltiples y diversos dispositivos IoT, si bien con una demanda de tráfico agregada muy baja. Por su parte, la configuración radio de RSI#2 and RSI#3 obedece más a unos requisitos de velocidades de transferencia y demanda agregada de tráfico mucho mayores, que se refleja en el despliegue de múltiples micro celdas (NR Cell #2 y #3 a modo de ejemplo en la Figura 2a) en otro canal frecuencial (RF carrier#2; que podría ser, p.ej., un canal de 40 MHz en la banda de 3.5 GHz) y con una configuración basada en una gNB-CU VNF en PoP#2 y múltiples gNB-DU VNFs en PoP#1 y #3 de cara a facilitar los mecanismos de coordinación entre celdas. A diferencia de RSI#1, las dos *slices* RSI#2 y RSI#3 se soportan mediante las mismas celdas (NR Cell #2 and #3) por lo que es necesario configurar la funcionalidad gNB asociada (i.e. gNB-CU#2, gNB-DU#3 y gNB-DU#4) con las garantías adecuadas de aislamiento y protección entre el tráfico eMBB y el tráfico MC en situaciones de alta carga.

A. Activación de una nueva RAN slice

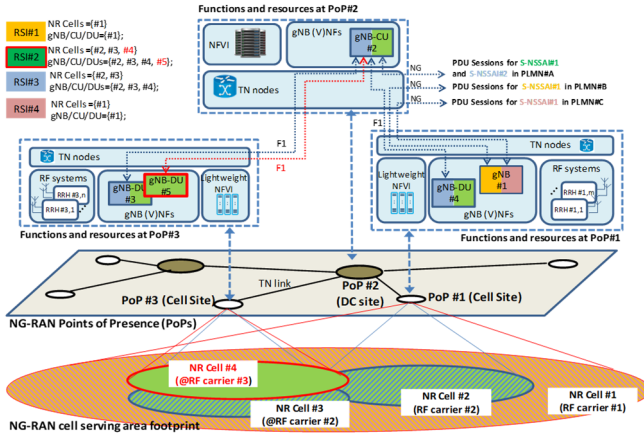
Partiendo de la configuración de referencia, se considera que surge la necesidad de proporcionar acceso a otro IoT SP especializado, por ejemplo, en el sector del coche conectado, y que pretende proporcionar un servicio de asistencia al conductor para seguridad vial que requiere poca capacidad agregada pero una muy baja latencia de entrega de mensajes. Se asume que dicho IoT SP cuenta también con su propia red 5GC privada (identificada como PLMN#C) y quiere ampliar la cobertura de sus servicios a la zona donde el gestor de RAN tiene desplegada infraestructura de acceso. Para ello, el gestor de la infraestructura RAN, a partir de una especificación funcional y de prestaciones del servicio (SLA) negociada con el IoT SP, utiliza las capacidades de planificación de red con soporte para modelado de RAN slices integradas en su NMS para determinar, por ejemplo, que la cobertura y capacidad proporcionada por NR Cell#1 resulta adecuada para la provisión del servicio si bien las necesidades de latencia de la nueva aplicación requieren que la celda NR Cell#1 opere también con una numerología (i.e. separación entre subportadoras) que permita tiempos de transmisión menores a los actualmente soportados. En base a ello, el NOP crea el correspondiente *RAN Slice Template* y mediante la interfaz de provisionado (véase Figura 1), procede a la creación de la nueva RAN slice (RSI#4), cuya configuración queda reflejada en la Figura 2b (las modificaciones respecto a la Figura 2a se resaltan en color rojo). Tal como puede observarse, la creación de RSI#4 básicamente conlleva en este caso una reconfiguración de las funciones gNB#1 para poder soportar la nueva numerología de baja latencia, provisionar los nuevos identificadores de red y *slice* (PLMN#C y S-NSSAI#1), y establecer el particionado de recursos pertinente entre RSI#1 y la nueva RSI#4. De esta forma, la celda NRCell#1 queda habilitada para proporcionar acceso a los terminales de ambos SP IoTs.



(a) Configuración de referencia



(b) Activación de una nueva RAN slice



(c) Modificación de la capacidad de acceso radio de RSI#2

Fig. 2. Ejemplos de despliegue y gestión de RAN slices

B. Incremento de capacidad de una RAN slice

Una vez activada la RSI#4, consideramos ahora que se presenta la necesidad de ampliar la capacidad de RSI#2 en una zona geográfica determinada y durante un periodo de tiempo determinado para cubrir una necesidad puntual de demanda del servicio eMBB proporcionado por el MNO. Al igual que en el caso anterior, la petición de incremento de capacidad conlleva una negociación previa entre el NOP y el MNO para formalizar los requisitos (SLA) seguida de un proceso de planificación de red por parte del NOP para determinar la implementación más factible. En este caso, se determina que la extensión de capacidad de la RSI#2 pasa por desplegar una nueva celda (NR Cell#4) en PoP#3 (por ejemplo, una celda en un nuevo canal frecuencial de 80 MHz en la banda de 3.7

GHz). El despliegue de dicha celda comporta en este caso la instanciación de un nuevo gNB-DU#5 en PoP#3 y su integración con el gNB-CU#2 existente. En base a ello, el NOP modifica el *RAN Slice Template* asociado a RSI#2 y procede a su modificación mediante la interfaz de provisionado. Los cambios realizados en RSI#2 se ilustran resaltados en rojo en la Figura 2c. En este caso, el procedimiento de provisionado comporta la actuación de las funciones MANO para la creación e interconexión del nuevo gNB-DU, la configuración y activación de la nueva NR Cell#4 y la posible reconfiguración de algunos parámetros radio en las celdas ya existentes (p.ej. lista de celdas vecinas en NRCell #2 y #3, ajustes de parámetros de movilidad entre celdas, etc.). Nótese que la automatización del proceso de provisionado es en gran medida dependiente del soporte de mecanismos SON que permitan cambios en la configuración de las celdas que no requieran de la intervención manual del NOP.

IV. CONCLUSIONES

El desarrollo de soluciones de gestión que permitan un aprovisionamiento dinámico y automatizado de *network slices* es fundamental para la consecución de una plataforma de acceso radio flexible y versátil que permita a los operadores hacer frente a la diversidad de usos previstos en escenarios 5G. A tal efecto, el marco funcional descrito en esta comunicación establece las diferentes funciones, modelos de información e interfaces que integrarían una solución de gestión de *slicing* para una infraestructura NG-RAN. La aplicabilidad del marco propuesto se ha ilustrado en un escenario de operador neutro mediante ejemplos que muestran los casos de activación y de modificación de *RAN slices* utilizadas para dar un servicio personalizado a aplicaciones y proveedores de servicio diversos.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la Agencia Estatal de Investigación (AEI) y el Fondo Europeo de Desarrollo Regional (FEDER) a través del proyecto SONAR 5G (ref. TEC2017-82651-R).

REFERENCIAS

- [1] P. Rost et al., "Network Slicing to Enable Scalability and Flexibility in 5G Mobile Networks," in *IEEE Communications Magazine*, vol. 55, no. 5, pp. 72-79, May 2017.
- [2] 3GPP TS 23.501 V15.0.0, "System Architecture for the 5G System; Stage 2 (Release 15)", December 2017.
- [3] 3GPP TS 38.300 V15.0.0, "NR; NR and NG-RAN Overall Description; Stage 2 (Release 15)", December 2017.
- [4] 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.
- [5] A. Ksentini and N. Nikaein, "Toward Enforcing Network Slicing on RAN: Flexibility and Resources Abstraction," in *IEEE Communications Magazine*, vol. 55, no. 6, pp. 102-108, 2017.
- [6] P. Caballero, A. Banchs, G. de Veciana and X. Costa-Pérez, "Multi-Tenant Radio Access Network Slicing: Statistical Multiplexing of Spatial Loads," in *IEEE/ACM Transactions on Networking*, vol. 25, no. 5, pp. 3044-3058, Oct. 2017.
- [7] O. Sallent, J. Perez-Romero, R. Ferrús, R. Agustí, "On Radio Access Network Slicing from a Radio Resource Management Perspective", *IEEE Wireless Communications*, October, 2017, pp. 166-174.
- [8] R. Ferrús, O. Sallent, J. Pérez-Romero and R. Agustí, "On 5G Radio Access Network Slicing: Radio Interface Protocol Features and

Configuration," in IEEE Communications Magazine, vol. PP, no. 99, pp. 2-10, Early access in IEEEExplore, January 2018.

- [9] 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.
- [10] 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.
- [11] I. da Silva et al., "Impact of network slicing on 5G Radio Access Networks," 2016 European Conference on Networks and Communications (EuCNC), Athens, 2016, pp. 153-157.
- [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)", January 2018.
- [15] 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. February 2018. Pre-print version available at arXiv:1803.01142v1 [cs.NI]