SPECIAL ISSUE ARTICLE

An adaptive 5G multiservice and multitenant radio access network architecture

Heinz Droste[1]*, Peter Rost[2], Mark Doll[3], Ignacio Berberana[4], Christian Mannweiler[2], Markus Breitbach[5], Albert Banchs[6] and Miguel A. Puente[7]

[1] Deutsche Telekom, Darmstadt, Germany
[2] Nokia Bell Labs, Munich, Germany
[3] Nokia Bell Labs, Stuttgart, Germany
[4] Telefonica, Madrid, Spain
[5] Deutsche Telekom, Bonn, Germany
[6] Universidad Carlos III, Madrid, Spain
[7] Atos, Madrid, Spain

ABSTRACT

This article provides an overview on objectives and first results of the Horizon 2020 project 5G NOvel Radio Multiservice adaptive network Architecture (5GNORMA). With 5G NORMA, leading players in the mobile ecosystem aim to underpin Europe's leadership position in 5G. The key objective of 5G NORMA is to develop a conceptually novel, adaptive and future-proof 5G mobile network architecture. This architecture will allow for adapting the network to a wide range of service specific requirements, resulting in novel service-aware and context-aware end-to-end function chaining. The technical approach is based on an innovative concept of adaptive (de)composition and allocation of mobile network functions based on end-user requirements and infrastructure capabilities. At the same time, cost savings and faster time to market are to be expected by joint deployment of logically separated multiservice and multitenant networks on common hardware and other physical resources making use of traffic multiplexing gains. In this context architectural enablers such as network function virtualization and software-defined mobile networking will play a key role for introducing the needed flexible resource assignment to logical networks and specific virtual network functions. Copyright © 2016 John Wiley & Sons, Ltd.

*Correspondence
Heinz Droste, Deutsche Telekom AG, Darmstadt, Germany.
E-mail: heinz.droste@telekom.de

1. INTRODUCTION

While the second generation of mobile communication systems focused on voice-centric services, the third and fourth generations primarily aimed at providing mobile broadband. In future mobile technology, this evolution will continue to facilitate a fully mobile and connected information society spanning a greater range of services, from human-centric to machine-centric communications. These services are characterized by a huge range of performance aspects such as high data rates, low latency, with some requiring ultra-high reliability, and high mobility as well as support of higher terminal density. These performance aspects are to be met while ensuring high levels of security, trust and privacy.

Future 5G networks are being designed to support this diverse set of complex performance requirements. Contrary to previous generations, 5G represents a shift in mind-set. It encompasses a holistic innovative network design to efficiently support applications with widely varying operational parameters, providing greater operational flexibility. As such, 5G is anticipated to be an important enabler of the future information society.

Former European Union funded R&D projects like Mobile and wireless communications Enablers for Twenty-twenty Information Society (METIS) [1] have laid important foundations for common understanding about what 5G could be. Three generic 5G services have been identified as the most important development drivers for mobile networks beyond 2020, that is, massive broadband, massive machine-type communication and critical machine-type communication.

Figure 1 illustrates the wide range of services and corresponding end-user requirements (see outer circle) as well
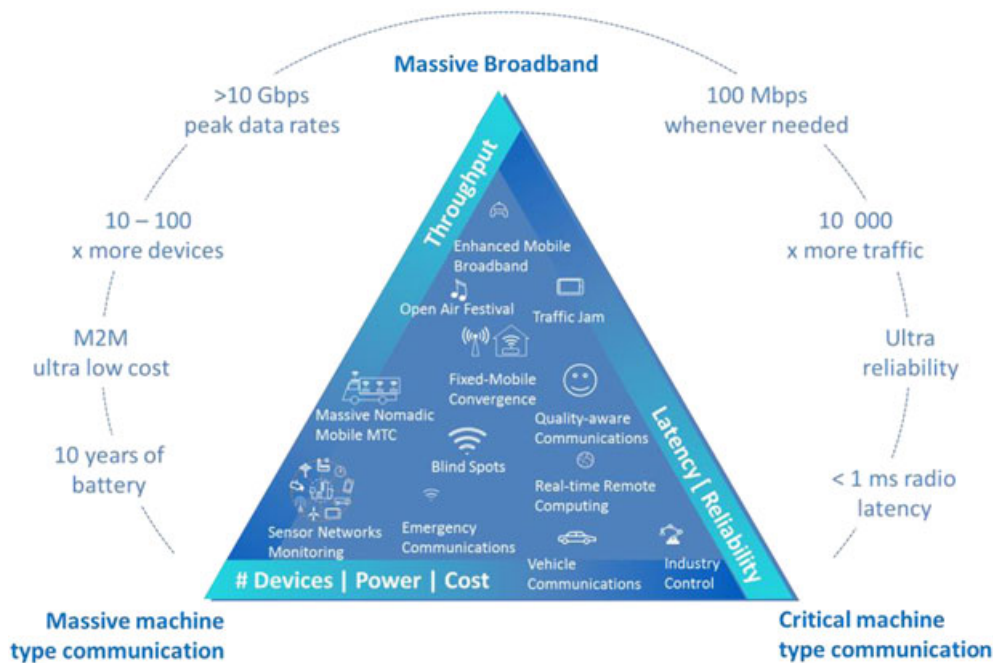
Figure 1. 5G NORMA use cases and respective end-user related requirements.

as end-user related use cases that are depicted within the triangle and described in more detail in [2].

Future services expect the network to connect to 10–100 times more devices per cell for massive machine-type communications, to provide peak data rates beyond 10 Gbps for massive broadband services and to enable ultra-reliable services with high availability and latencies of approximately 1 ms for critical machine-type communication. Specific performance requirements are driven by different use cases. Devices for sensor networks must be low cost and energy efficient, whereas devices for vehicular communications must ensure that they can communicate even if there is no network coverage available. This extreme range of partly contradicting requirements cannot be realized by specific network components at the same time. Hence, future networks need to be much more flexible in order to be re-configured to meet these apparently contradictory requirements. The challenge is to meet the diverse set of requirements with a network able to support diverse operational needs and services, with the scale economies of one network.

The remainder of this article is structured as follows: Section 2 describes the main objectives and the overall concept of 5G NORMA. Outcomes of the first half year's activities have yielded a reference architecture, which can be evaluated from four different perspectives, and are presented in Section 3. In order to ensure that project outcomes are mature and can be implemented in subsequent phases, the project is structured around being able to demonstrate comprehensive verification and validation against the conceptual objectives. This is described in Section 4. Conclusions and next steps are summarized in Section 5.

2. OVERALL CONCEPT

2.1. Key enablers

The main goal of 5G NORMA is to propose a multiservice mobile network architecture that adapts the use of the mobile core network (CN) and radio access network (RAN) resources to the service requirements, the variations of the traffic demands over time and location and the network topology including the available fronthaul and backhaul capacity. An approach for achieving the previous objective is to decompose the mobile network functions including access and core functions and adaptively allocate them within the network topology, depending on the transport network capabilities, computing resources and specific services and their requirements, for example, bandwidth and latency. This adaptive allocation of functions incorporates several advantages:

- If service requirements and backhaul capacity allow for centralizing the functionality in the network cloud, better scalability and pooling gains can be obtained from moving all the functionality to central clouds.
- If services have special requirements, for example, extremely low latencies, which require moving part of the functionality to the access, or backhaul constraints do not allow full centralization, a fully or partially distributed deployment of network functionality provides better performance.

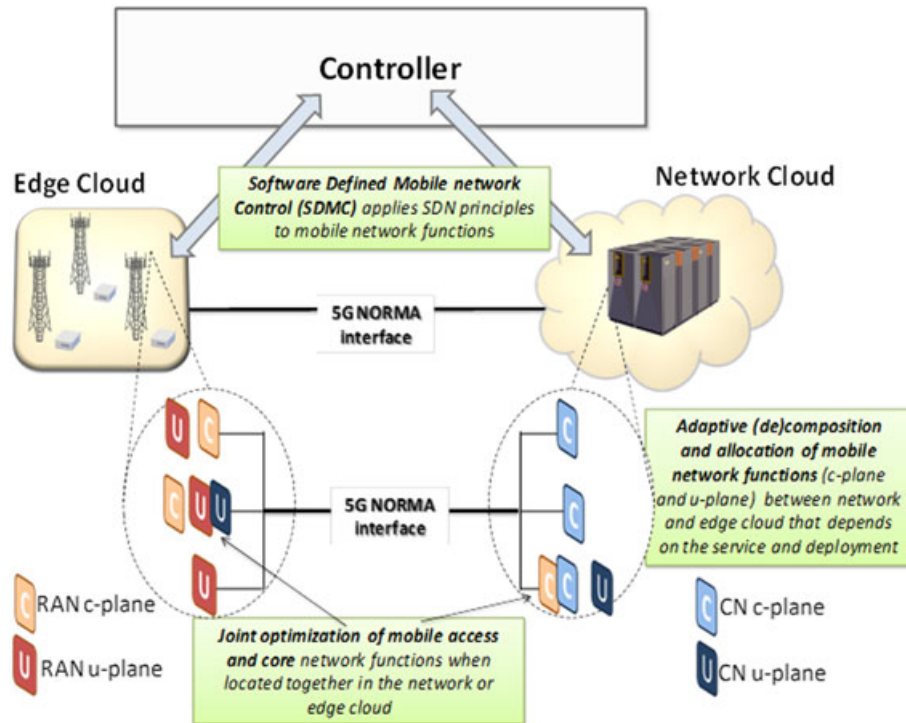Figure 2 shows the fundamental components of the 5G NORMA architecture:

Figure 2. 5G NORMA concept based on three innovative enablers.

- an edge cloud, which is the integration of radio base stations and distributed servers at the radio or at the aggregation sites;
- a (central) network cloud, which consists of servers at central sites; and
- controllers, which are responsible for controlling all functions executed within the network according to the software-defined mobile networking approach, which is explained in Section 3.1.

Along with these fundamental components, the figure also illustrates the three enabling technologies of 5G NORMA:

- an adaptive (de)composition and allocation of mobile network functions between the edge and the network cloud depending on the service requirements and deployment needs,
- software-defined mobile network control (SDMC), which applies the software-defined networking (SDN) principles to mobile network specific functions; and
- joint optimization of RAN and CN functions localized together in the network cloud or the edge cloud. Figure 2 indicates exemplary at the 5G NORMA interface three services that may use different allocation of RAN and CN functions: while for some services, most of the functionality may be located in the edge cloud; for others, it may be located in the network cloud. It can also be seen that a distinction is made between control plane functions that handle

signalling traffic and user plane functions that process user traffic.

Based on principles of standardization at ETSI on network function virtualization [3] and integration of SDN [4], 5G NORMA develops a virtualization of mobile network functions, which enables a flexible decomposition of mobile network functions between the radio access and the network cloud infrastructures, supporting fully distributed, partially distributed, and fully centralized deployments. The decision about the degree of centralization, that is, allocation of network functions, takes into account service and application requirements driven by the optimization of metrics such as latency, QoE,[†] resource utilization and energy efficiency. As explained in more detail in Section 3.1, adaption to service requirements is carried out by instantiation of different network slices where each slice is tailored to the specific requirements of the corresponding service. In addition to service requirements, the allocation of network functions also takes into account the computational requirements, existing or required backhaul deployments and associated transport capacities and the spatial and temporal characteristics of traffic within the RAN.

For instance, we may choose different configuration settings as a function of the backhaul characteristics, that is, the transport network connection to the antenna site, for

---

[†]QoE, Quality of Experience.

example, a fully distributed configuration, a fully centralized one and an intermediate one.

Current SDN implementations focus mostly on wired networks to separate routing control (control-plane) from routing execution (user plane). 5G NORMA applies these principles to wireless functions beyond routing, where benefits of this technology may be even more significant than for wired networks, as the control functionality of wireless networks includes many more wireless related functions than just routing control. This includes time critical functions such as scheduling control, modulation and coding scheme selection and HARQ processing and other less time critical functions such as radio resource control, power control, handover decision and execution. With the SDMC concept, all these functions can be implemented more easily by a programmable and logically centralized control, which provides very important benefits for the flexible operation of the wireless edge network.

Today, the physical separation between RAN and CN functions limits the interaction between those two logically separated domains requiring the specification of complicated interfaces, which in turn delays or impedes innovation.

The adaptive (de)composition and allocation of mobile network functions described earlier may cause that the border between RAN and CN becomes more blurred. This provides new opportunities to jointly optimize the operations of RAN-specific and CN-specific network functions, which previously were dealt with separately because of their different locations.

2.2. Novel functionalities

Although current network architectures and concepts already support a variety of service classes and corresponding QoE/QoS,[‡] network functions typically can neither be reallocated nor adapted to service classes. Current architectures are based on a pre-defined set of bearers, support only limited variations of the topology and, thus, prevent the networks from efficiently supporting new and dynamically changing service classes.

The previusly described enabling technologies allow for introducing logically separated network slices, which build the basis for novel functionalities such as multiservice and multitenancy. A network slice in that sense is a collection of (mobile) network function instances including their required resources necessary to operate an end-to-end (self-contained) logical mobile network.

Supporting both functionalities signalling-based dynamic network reconfiguration will enable creation of slice-specific service chains that fulfil the requirements of offered services as well as tenant-specific Service Level Agreements (SLAs), respectively.

[‡]QoS, Quality of Service.

3. REFERENCE ARCHITECTURE AND DIFFERENT VIEWS

In the past, the architecture of legacy networks has typically been described by defining functional blocks and logical interfaces in between. As 5G NORMA aims at introducing flexible and adaptive allocation of mobile network functions, this pure functional view is not sufficient anymore. In the following sections, four different views on the initial reference architecture are described. More details can be found in [5]. During the course of the project, it is intended to have a number of iterations in order to refine the functional architecture and detail functional blocks and network functions as well as interface definitions.

3.1. Functional view

The functional view captures the functional blocks and the logical interfaces regardless of each function block's location within the network topology and regardless of the resources used. From functional point of view (cf. Figure 3), we distinguish

- service layer,
- management and orchestration layer, and
- control and data layer.

The service layer is accessible to the tenants and includes information on available services as well as business support systems and policies. The management and orchestration layer encloses functional blocks that contain network management and orchestration (MANO) functions [6]. At this layer, slices are created by collaboration of the service, inter-slice and slice orchestrators. Virtual network functions (VNFs) are instantiated by the VNF managers, and overall network management is carried out by the operation support system and element managers, which are already present in today's non-virtualized networks. In this context, it is important to note that provisioning a timely status of the whole network is supported by vendor-specific virtual infrastructure managers (VIM) that are assisted by locally responsible VIM agents.

The control and data layers (user plane) in this view give information on the arrangement of the VNFs and, in addition, bare metal functions called Physical Network Functions (PNFs). These PNFs take over computation intensive and time critical tasks and are mainly located in the vicinity of the air interface in the RAN.

5G NORMA distinguishes functions that are being 'controlled' and those functions that 'control' the network. Functions in the data layer, that is, functions that are being controlled, comprise typical user plane functions. But there exist also functions in the control layer, which process signalling traffic that are being controlled. The most important function in the control layer is the SDM controller, which controls the overall network including all PNFs and VNFs in data and control layer. In difference to the MANO
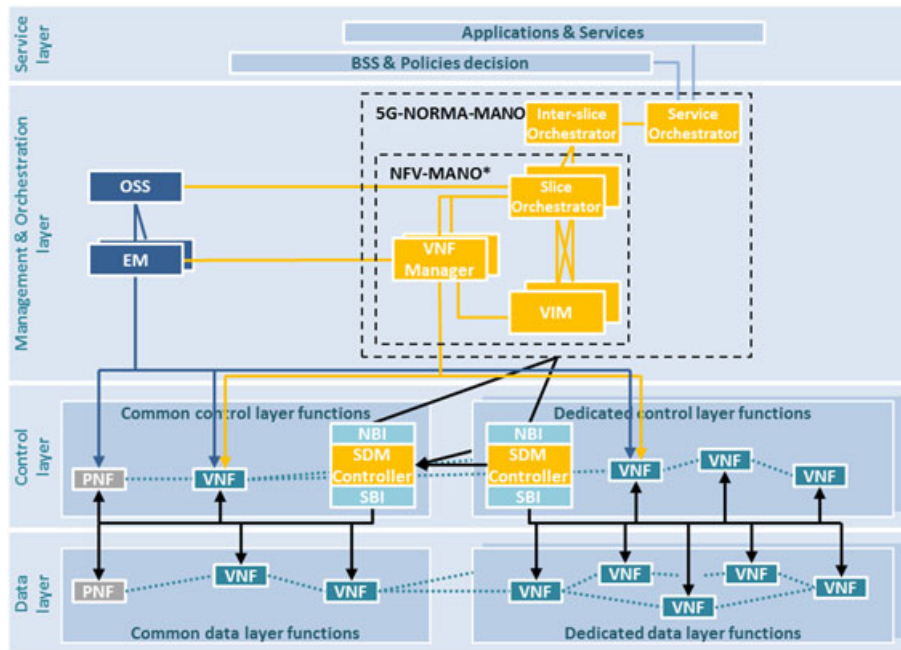
Figure 3. Preliminary 5G NORMA functional reference architecture. OSS, operation support system; EM, element managers; VNF, virtual network functions; VIM, virtual infrastructure managers; PNF, Physical Network Function; SDM, software-defined mobile networking; MANO, management and orchestration.

functions, the SDMC is working on a time basis that is very close to the radio frame structure and, hence, there is a capability to control scheduling, interference management and other radio related functions synchronously to the radio frames.

Control layer functions as well as data layer functions can be either common to multiple network slices or dedicated to a single network slice. Functions assigned to multiple logical network slices are often related to physical resources. For example, as radio spectrum resources cannot be easily increased on demand, it may be beneficial to pool these resources for multiple slices and control their usage by a function that is common to these slices. Common and dedicated VNFs are controlled by distinct SDMC entities: The owner of a network slice is given full control over its own slice through the dedicated SDMC, but aspects affecting other network slices are covered by the common SDMC to remain under control of a 'neutral party'.

The 'southbound' interfaces between the controllers and the VNFs and PNFs must be standardized and supported by all deployed functions. The interface in addition must provide sufficient flexibility to programme the desired network behaviour. By means of customized control applications, SDMC effectively provides such network programmability capabilities for Mobile Network Operators (MNOs) as well as third parties, such as, virtual operators or vertical industries. Further, multitenancy mobile networks inherently introduce the notion of shared network functions. Therefore, 5G NORMA develops the concept of SDM coordination, which

extends SDMC towards a multitenant control of shared network functions.

In the management and orchestration layer, function blocks are shown concretely and are present for all network slices. By contrast in control and data layer, the VNFs and PNFs are shown in an abstract way because each network slice may run a different set of VNFs and PNFs, specifically adapted to the service, which is implemented by the network slice and to the physical deployment on which it is executed. Nonetheless, defining and detailing these control and data layer functions will be matter of further work in 5G NORMA. The blocks in the functional view refer to logical elements and not to virtual or physical resources (Section 3.2).

3.2. Resource view

The resource view comprises a comprehensive set of resources available to network management and orchestration entities in order to compose mobile network instances for different use cases and tenants. First of all, it is distinguished between the different deployment types which denote the data centres at central network and edge cloud nodes as well as the bare metal nodes in the vicinity of antennas where PNFs are closely related to specific hardware (Figure 4). There are only a few central network clouds in an operator nationwide network, whereas the number of edge clouds is at least an order of magnitude higher than the number of central clouds. Edge clouds are expected to be deployed in rather densely populated areas
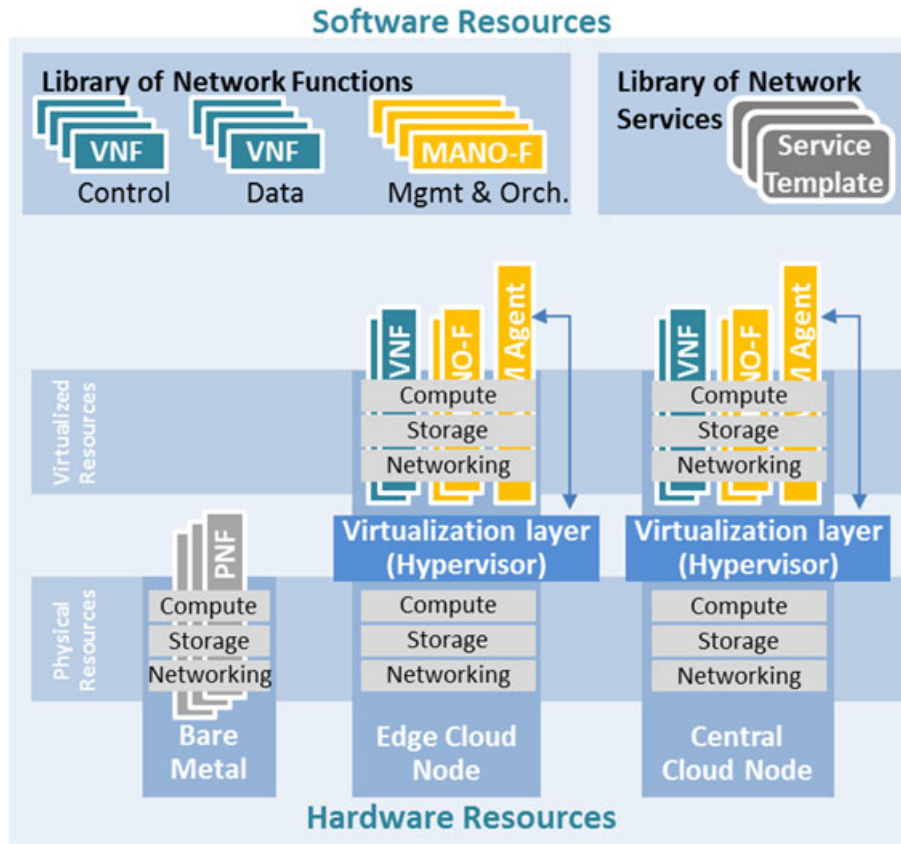
Figure 4. 5G NORMA resource view. PNF, Physical Network Functions; VNF, virtual network functions; MANO, management and orchestration.

or directly at the radio sites. The cloud associated resources at the data centres (compute, storage, and networking) can be based on commodity hardware.

The second part of infrastructure resources considered are software resources such as network function and service template libraries. The library of NFs represents the repository of all executable VNF packages including the necessary blueprint and metadata, for example, resource requirements, supported interfaces and reference points as well as orchestration and configuration parameters. The library of network services represents the repository of all available network services including the necessary blueprints and metadata such as QoS parameters. A network service template refers to the set of VNFs and PNFs that should be chained to implement the network service, e.g. VoIP or IMS.

As can be seen in Figure 4, virtual and physical resources for compute, storage, and networking have to be assigned to VNFs and MANO functions at central as well as at edge clouds, whereas those resources at bare-metal nodes are statically assigned to PNFs. Virtual and physical resources are abstracted from each other by a virtualization layer (hypervisors). On request of the orchestrator framework, hypervisors take the role of assigning specific amount of resources (compute, storage and net-

working) to VNFs. Becasue of traffic fluctuations, VNFs may need to be rescaled from time to time, that is, adapting the resource assignment properly to changing demands. In combination with the deployment and topology view, which is described in the following sections, the resource view plays an important role for further development of the 5G NORMA architecture especially with respect to economic evaluations.

In order to enable different business models, the interfaces between physical, virtual and MANO functions have to be designed properly from the beginning throughout the process of the architecture definition. For example, a tenant could rent an end-to-end network slice from a service provider owning the software resources (Libraries, MANO-F and VNFs). The service provider on his part relies on the infrastructure of a virtual infrastructure provider that owns virtual and physical compute, storage and networking resources as well as bare metal nodes.

3.3. Deployment view

The deployment view shows the abstract mapping of functional blocks to different resource classes. In this context, abstract mapping implies that the opportunities for meaningful placement of functional blocks are depicted instead of concrete instances.
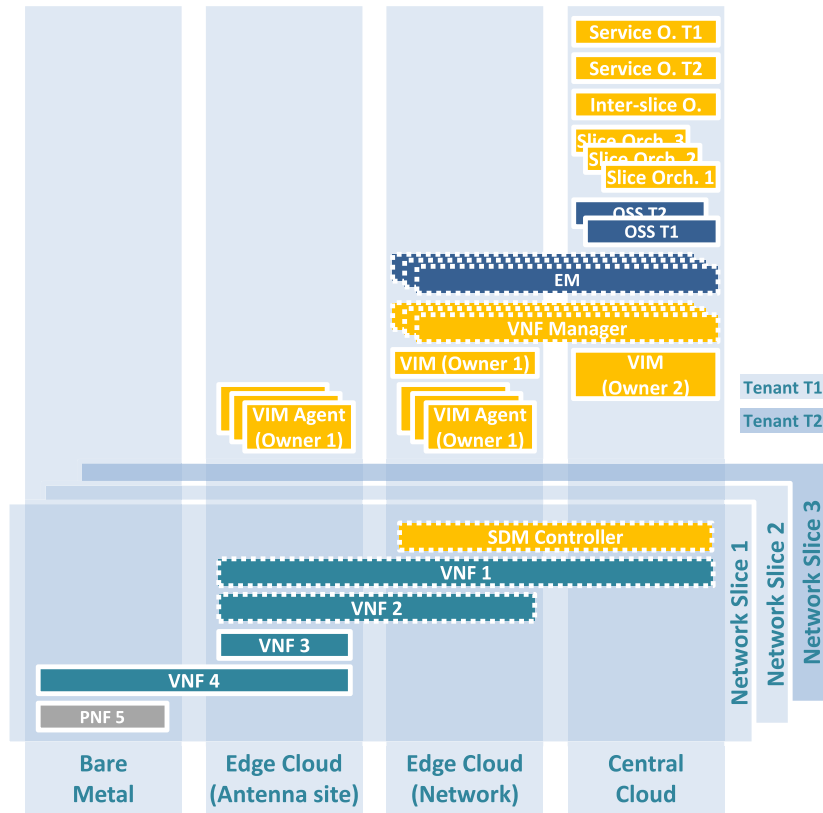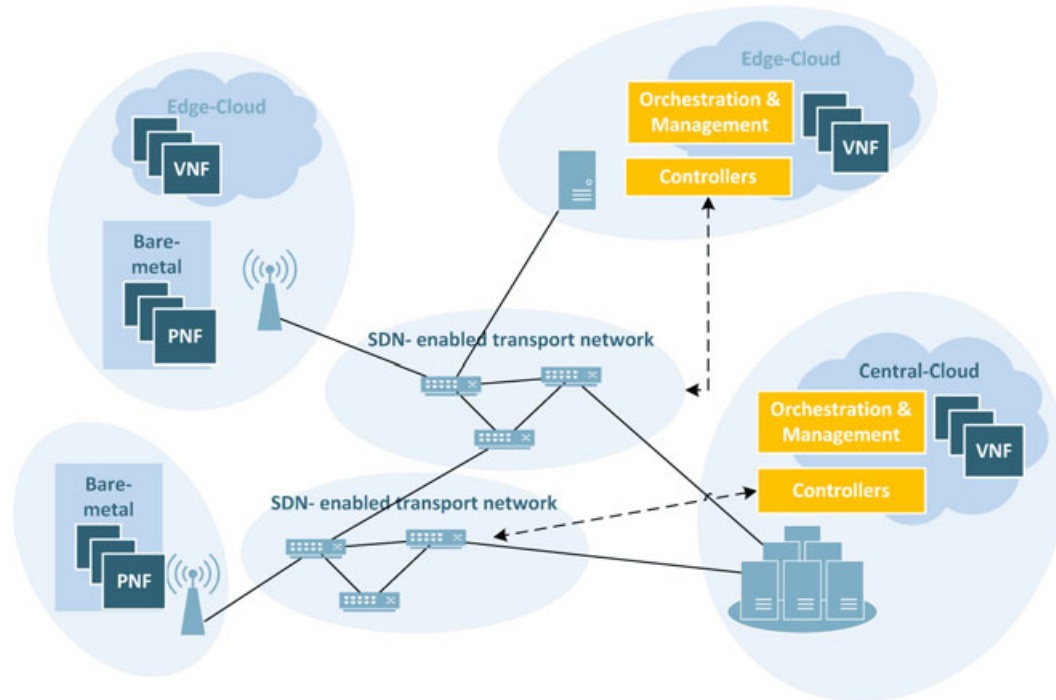
Figure 6. 5G NORMA topology view. VNF, virtual network functions; VIM, virtual infrastructure managers.

between an edge cloud data centre and the radio access point at the antenna site are low latency and high capacity. In the case of centralizing all radio access protocol layers, latency should be less than $100-200\ \mu s$ [7, 8]; that is, the distance between edge cloud and antenna site should not exceed a few 10 km, and a dedicated point-to-point connection should be used. In case the connection between antenna site and edge cloud cannot support such low latencies, 5G NORMA provides sufficient flexibility to deploy a different functional placement with only partly centralized layers, as far as permitted by the latency requirements of the implemented service.

The required capacity of the fronthaul connection depends on the implemented functional split, the properties of the radio signal that shall be transmitted over the air interface, in particular on the signal bandwidth and on the number of antennas. It is typically in the range of $1-10$ Gbps per antenna. Redundancy is usually not required, as a link failure will affect only few cells and thus a limited number of terminals. Aside the centralization of major parts of the radio access protocol processing, also dedicated base stations, either macro or small cells, can be installed at the antenna site. The backhaul of such base stations is usually based on optical fibres. Micro wave links are cheaper to build than optical fibres, but the achievable capacity of micro wave links is significantly lower.

It is important to note that the opportunities for placement of functional blocks at the different nodes depend heavily on the virtual network and front haul capabilities in the network topology.

4. USE CASE VALIDATION AND ARCHITECTURE DESIGN VERIFICATION

It is important that project outcomes are mature and can be handed over to next step realization activities. Therefore, the fulfilment of end-user requirements is checked by a comprehensive use case validation, and the fulfilment of technical and economical requirements is checked by architecture design verification. For use case validation, 5G NORMA has defined 12 use cases that have been analysed in order to identify relevant functional and performance requirements. In order to validate the fulfilment of the identified requirements by 5G NORMA enhanced functionalities, three different scenarios have been also defined. These scenarios consist of two or more use cases that consider different services and assumptions. Hence, each scenario involves different requirements in order to test the expected flexibility, which needs to be supported by the network architecture. Use cases and scenarios together with Key Performance Indicators (KPIs) and other evaluation criteria (Figure 7) provide the basis for checking the fulfilment of end-user requirements as well as system and economic requirements.

The objective of architecture design verification is guiding a two-step architecture design iteration and thereby to finally provide a proof of concept of the 5G NORMA key innovations. The assessment will be based on evaluation metrics to be elaborated based on the use case and scenario definitions in [2]. For architecture design verification, architecture options will be defined that include different possibilities for placement of management and orchestra-

Figure 7. Evaluation criteria.

tion functions, alternative structures for transport networks (backhaul and fronthaul) or alternative instantiations of functions chaining within logical slices.

The most important evaluation criteria are depicted in Figure 7. Fulfilment of performance requirements defined by the use cases will be checked by system level simulations of the scenarios defined. Demonstrations will prove the feasibility of network reconfiguration. A qualitative protocol analysis planned in the course of the project will check the fulfilment of functional requirements, while a protocol overhead analysis will evaluate the benefit of adaptive (re)allocation of network functions and SDN-based mobility management. Operational requirements will be mainly defined and checked by operator expertise [9]. Security risks and vulnerabilities will be checked from an architectural point of view.

An important contribution to the final proof of concept is economic evaluations that will make sure that, on one hand, multiservice and multitenancy in connection with network virtualization (instead of operation of separated networks) will bring the expected cost benefit, while on the other hand, it will also provide methodologies for estimating future 5G business cases and identifying new business roles. Last but not least, soft KPIs such as feasibility, complexity and required standardization effort, as well as the evaluation of scalability of network and MANO functions under real network conditions, will contribute to the final assessment showing that innovations introduced by 5G NORMA hold not only for demonstrator conditions but also in real operator networks' operational conditions.

5. CONCLUSIONS AND NEXT STEPS

The architectural flexibility of 5G NORMA supports (de)composition, adaptation, joint optimization, placement and software-defined control of mobile radio access and core network functions in a way that is specifically tailored to the service and deployment needs, thereby providing a flexible multi-service and multitenant operation on shared infrastructure.

As already mentioned, next steps include defining and detailing concrete control and data layer functions as well as their control under the SDMC paradigm. One of the most challenging tasks will be to further

elaborate on the common network functions including their control (referred to by Software Defined Mobile Coordinator (SDMX)), because here services and tenants that are considered fully independent are sharing functionality for efficiency reasons. Finally, the "5G NORMA interface" (cf. Figure 2), namely, its possible manifestations as data and control layer interface, respectively, and the interfacing towards the MANO layer and SDMC need to be worked out with the challenge of finding the right balance between flexibility and complexity.

ACKNOWLEDGEMENTS

REFERENCES

1. METIS D6.6 "Final Report on the METIS 5G System Concept and Technology Roadmap" ICT-317669 METIS Deliverable 6.6. Version 1, 2015.

2. EU H2020 5G NORMA. D2.1: Use cases, scenarios, and requirements, 2015.

3. ETSI Industry Network Functions Virtualisation Industry Specification Group (NFV ISG) white papers, specificat ions, etc. Available from: http://www.etsi.org/technologies-clusters/technologies/nfv [accessed on 26 July 2016].

4. Software-Defined Networking. the new form for networks, ONF white paper. Available from: https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/wp-sdn-newnorm.pdf [accessed on 13 April 2012].

5. EU H2020 5G NORMA. D3.1: functional network architecture and security requirements, 2015.

6. ETSI GS NFV-MAN 001. Network functions virtualisation (NFV); management and orchestration, 2014. V1.1.1 (2014-12).

7. Smith PA. CPRI 'FrontHaul' requirements discussion with TSN San Diego, CA, 2014. Available from: http://www.ieee802.org/1/files/public/docs2014/new-ashwood-tsn-cpri-fronthaul-0714-v03.pdf [accessed on 26 July 2016].

8. Son HJ, Shin SM. Fronthaul size: calculation of maximum distance between RRH (at cell site) and BBU (at CO). Netmanias tech-blog. Available from: http://de.slideshare.net/Netmanias/netmanias20140401 calculation-of-fronthaul-fiber-distance-en [accessed on 1 April 2014].

9. NGMN Alliance: 5G White Paper. Version 1.0. Available from: https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf [accessed on 17 February 2015].