



HAL
open science

Complexity in 5G Network Applications and use cases

Rui Li, Bertrand Decocq, Anne Barros, Yiping Fang, Zhiguo Zeng

► **To cite this version:**

Rui Li, Bertrand Decocq, Anne Barros, Yiping Fang, Zhiguo Zeng. Complexity in 5G Network Applications and use cases. 31st European Safety and Reliability Conference, Sep 2021, Angers, France. pp.3054-3061, 10.3850/978-981-18-2016-8_211-cd . hal-03363854v1

HAL Id: hal-03363854

<https://hal.science/hal-03363854v1>

Submitted on 4 Oct 2021 (v1), last revised 28 Jan 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COMPLEXITY IN 5G NETWORK APPLICATIONS AND USE CASES

Rui Li, Bertrand Decocq

Orange Innovation, Châtillon, France. E-mail: rui.li@orange.com, bertrand.decocq@orange.com

Anne Barros, Yiping Fang, Zhiguo Zeng

Laboratoire Génie Industriel, CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette, France.

E-mail: anne.barros@centralesupelec.fr, yiping.fang@centralesupelec.fr, zhiguo.zeng@centralesupelec.fr

The fifth generation (5G) of mobile telecommunication network is designed with an ambition to be a network faster, stronger, better and smarter than its predecessor. With the digital transformation, all industry sectors will develop new applications with new requirements regarding telecommunication networks that 5G should be able to meet. To meet the requirement of future 5G use cases and applications, it is crucial to study the complexity of such network system by distinguishing different parts, layers, components as well as their interdependencies. This paper describes the 5G networks from an End-to-End perspective (device, radio network, core network, data network) and from a multi-layer perspective (orchestration, virtualisation/containerization and infrastructure) to show how this system (or system of systems) is complex, especially when we address resilience challenges. Resilience requirements and challenges are further explained by proposing relevant scenarios and use cases. In this paper, we mainly intend to highlight 5G network complexity and open a discussion on methodologies to model such complex network for its resilience study with the hope that this paper could inspire the future study of researchers in the related field.

Keywords: 5G Network, resilience quantification, resilience metrics, network applications, complex system, vertical requirements.

1. Introduction

The telecommunication domain keeps evolving rapidly since its birth. From the first generation cellular network to the newest generation, every one of them brings convenience to daily life and work. As the key to the future technology ecosystem, the fifth generation (5G) of mobile telecommunication network is designed to be a faster, stronger, better and smarter telecommunication network than ever before.

By upgrading existing technologies and incorporating new technologies, 5G networks are without a doubt a promising solution for future telecommunication needs.

In the Radio Access Network (RAN), new radio technologies are applied to 5G networks. Terminals are eligible to use both 5G and 4G frequency bands and connect to both 4G and 5G antenna. The utilization of orthogonal frequency-division multiplexing (OFDM) allows multiple communication channels to coexist, and thus it is possible to treat high frequency and low frequency bands at the same time to obtain both higher bandwidth and wider coverage. Intelligent antennas using “massive MIMO” are implemented which can further improve network capacity (IEEE 802.11ad, 2012; Patriciello et al., 2020). Besides, the beamforming technology ensures to transmit signal in a specific direction where it is useful to users rather than sending in all directions, such that less interference is created and less energy is consumed.

The 5G core network (5GC) becomes a service based architecture. In this software designed architecture, each Network Function (NF) is delivering “services” to other NFs to access control plane functionalities, subscriber or network data repositories through an interface of a common

framework (3GPP TS 23.501, 2021; Mademann, 2018). To deliver services more dynamically, 5G networks adopt the techniques such as Network Function Virtualisation (NFV) and Software Defined Networking (SDN).

By introducing the concept of slicing into the network, it is possible to create different virtual networks for different services. In 5G, such network will be a dedicated slice providing tailored network capabilities and network characteristics according to the requirement from the customers by respecting specific rules without disturbing the rest of the network outside the slice. Multiple users, if permitted, can connect to one same slice. One user equipment, if needed, may have access to multiple slices at one time. A network slice subnet represents a group of network functions that form part or complete constituents of a network slice. A network slice subnet may contain for example instances of Core Network functions only, or instances of Access Network functions only, or any combination thereof (3GPP TS 28.530, 2021).

Edge computing is a generic term encompassing a variety of different approaches to put computing and storage resources at the edge of the network close to the customer rather than in remote datacenters. Initially, this notion was introduced and used for mobile networks, hence the term Mobile Edge Computing (5G Smart, 2020). Later, the European Telecommunications Standards Institute (ETSI) defined the term Multi-access Edge Computing (MEC) as a generalization of Mobile Edge Computing to any network (ETSI, 2019). Some latency-sensitive network application functions can be deployed on MEC servers near the RAN or even at the macro base station. Therefore, some of the data will be stored and proceeded in distributed edge cloud services.

From an End-to-End perspective, these technologies break the boundary of different parts of network resources. From a multi-layer perspective, with NFV, a virtualization layer is added into the network architecture. Thus, 5G becomes a complex system and even a system of systems.

Combining the afore-cited technologies, 5G networks are going to greatly reshape the domain not only by its performance but also by offering a transition from a “horizontal” service delivery model toward a “vertical” service delivery model (Banchs et al., 2019). The former model provides identical services to all customers, while the latter provides tailored and personalized services for certain industry sectors. Such “vertical” delivery model introduces new scenarios and application use cases to 5G networks.

International Telecommunication Union (ITU-R, 2015) classifies 5G services into three categories: enhanced mobile broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable and low-latency communications (URLLC). Virtual or Augmented Reality, one of eMBB use cases, which constantly transfers a huge amount of data could benefit from 5G high speed network connection. In the example of smart fabrication, a massive machine-type communications service, the network slicing may help the factory to check its production by connecting IoT equipment to a dedicated slice. Autonomous-driving, as a typical URLLC service, its massive data can be calculated in real time at the edge with the help of 5G networks.

With the digital and technological transformation, 5G can be applied to more and more scenarios, nevertheless, it also creates a resilience challenge not only for the service providers but also for all the supported industries or verticals. The more technologies are integrated into the system, the more potential risks. Furthermore, 5G becomes a complex system, one single failure, if not fully fixed, could propagate from one single point to a series elements of the system. Thus, it is crucial to study the complexity of such network system by comprehending each part and each layer of it. Only when we have a full understanding on how 5G networks are composed and how a Network Service is established, can we analyze the risk and resilience of the system.

The paper is structured as follows: in Section 2 we explain 5G complexity from an End-to-End perspective by decomposing each part of the End-to-End service; Section 3 is devoted to presenting a multilayer perspective of 5G network, where the Network Function Virtualisation is mainly discussed; in Section 4 we briefly describe the process of setting up a service using the technologies that we introduce; Section 5 takes on the resilience challenge by introducing adverse events in telecommunication network and recent works on performance evaluation in 5G system; Finally we conclude the paper by providing suggestions for further research in Section 6.

2. The complexity from the End-to-End perspective

Figure 1 shows the End-to-End architecture in 5G network. It includes in general terminals, Next generation Radio

Access Network (Ng-RAN), Transport Network, 5G Core Network, Data network. In a 5G use cases such as autonomous vehicles, more than one Network Service may be needed. A Network Service may traverse all or only part of the aforementioned elements.

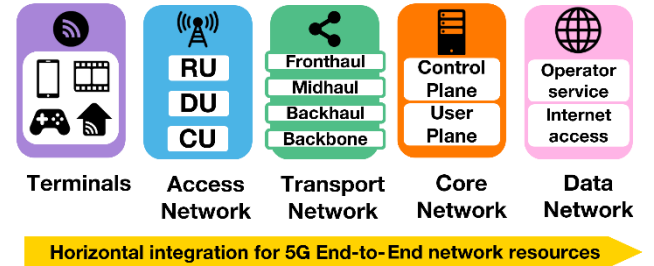


Fig. 1. 5G End-to-End Architecture.

2.1. Terminals

A Network Service normally starts from an end-user device. These devices are called terminals or user equipment (UE). Since 5G will be widely applied in telecommunication, variant terminals will be connected to the network. These terminals could be smart phones, vehicles, IoT terminals, etc. Among them, some terminals require simultaneous connections to multiple services. More specifically, a typical application would be the case where a smartphone plays an online football match for the user while in the background the device tries to get a push notification from mail service. In other cases, some devices may require the same service at the same time. Such scenario commonly happens in a factory. When an environment change is detected, all connected sensors will report this abnormal data or the extracted information to the central server at the same time. In the multiple services situation, the isolation between Network Services must be guaranteed to meet the requirement of each one of the services. In the massive devices access situation, the system must be resilient enough to cope with a potential congestion in the data processing and transport.

2.2. Radio access network

For a wireless terminal, to transfer data from the UE to the Radio Access Network (RAN), the data will be firstly received by the antenna embedded in Radio Unit (RU). Baseband Unit (BBU) connected to the RU will then transport a baseband frequency before sending the data to the edge or Core Network. Open RAN as a future generation of RAN chosen by 5G network, provides a standardized interface between RU and BBU as well as a standardized interface between BBUs to cooperate with multiple vendors. With virtualized RAN (vRAN), the BBU can be virtualized on multiple NFV platforms and be shared with operators (see Section 3.1 for more details on Network Function Virtualisation). BBU can be divided into multiple parts. The first part, distributed unit (DU), takes charge of real-time BBU scheduling functions, while the second part, centralized unit (CU), completes the non-real-time BBU functions. Some software parts of BBU will be placed

together with RU. CU and DU can be deployed flexibly, namely co-located with RU, in edge cloud or regional datacenter. Virtualisation and standardized interface make the most of open interfaces by enabling sharing CU and DU with multiple vendors (Wind River, 2017).

2.3. Transport network

To ensure a highly reliable and good performance network, Transport Network plays a crucial role. Transport Network includes the fronthaul of remote units, the backhaul of base stations, optionally a midhaul between the distributed and centralized units, and the backbone between core datacenters. Different transmission technologies are used for each part of Transport Network, for example: dark fiber for fronthaul and midhaul with direct connections between the nodes (RU to DU and DU to CU respectively), WDM rings for backhaul and backbone networks. Network Slicing can be based in the first step on VLAN/VPN for each transport segment (called basic soft-slicing or logical isolation between slices), and later on new technologies like Segment Routing-Traffic Engineering (SR-TE) for enhanced soft slicing with specific performance or designed per type of slice, and in the third step on Flexible Ethernet (FlexE) or Time Sensitive Networking (TSN) for hard slicing where the slices are fully isolated with guaranteed services performance. For resilience purposes, the IP network, from the Edge to the Core Network, is doubled and relies on WDM rings, and is able to react in 50 milliseconds in case of failure.

2.4. Core network

In 5GC, one of the most important characteristics is the separation of the User Plane (UP) functions from the Control Plane (CP) functions (3GPP TS 23.501, 2021). UP functions mainly take care of traffics forwarding while the CP functions manage the authentication, network slice selections, etc. The principal advantage of such separation is being able to flexibly scale the CP functions independently on UP functions in case of traffic peak vice versa. Another benefit lies in the flexibility to separately deploy CP functions so that some functions can be deployed, according to the requirement of the use case, in a centralized datacenter or a distributed one close to the RAN. The flexibility in scaling and deployment completely makes 5G networks more complex than the last generation.

Core functions

Figure 2 depicts a 5G network architecture. The upper part of the architecture shows the 5GC Control Plane which uses Service-based interfaces. The 5GC Control Plane consists of the following Core Network Functions (NF).

- Authentication Server Function (AUSF).
- Access and Mobility Management Function (AMF).
- Data Network (DN), e.g. operator services, Internet access or 3rd party services.
- Network Exposure Function (NEF).
- Network Repository Function (NRF).

- Network Slice Specific Authentication and Authorization Function (NSSAAF).
- Network Slice Selection Function (NSSF).
- Policy Control Function (PCF).
- Session Management Function (SMF).
- Unified Data Management (UDM).
- Unified Data Repository Function (UDF).
- Application Function (AF).

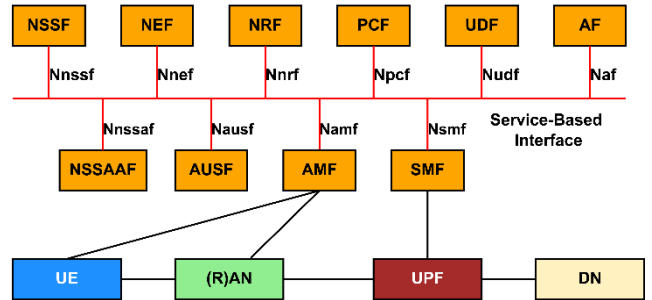


Fig. 2. 5G Service-Based Architecture.

Mobile Signaling is the engine of Mobile Networks Control Plane. As the Core Network adopts a service-based architecture, the signaling network elements are designed in form of Control Plane Network Functions. Each NF can thus expose a set of services called NF services to the service-based interface. Thus, these NFs can be both a consumer where they seek to consume the NF services provided by other NFs and an NF service producer where they provide their exposed services to NF service consumers. Each NF can provide multiple NF services for different NF consumers and can consume NF services from multiple service providers.

To avoid ambiguity in the paper, NF service refers to part of functionality of a Network Function that can be consumed by other NF, while Network Service defines a set of NFs connected together that facilitate a network operation.

2.5. Edge computing

Edge computing is an optional solution for 5G networks. The presence of Multi-access Edge Computing (MEC) reduces some Network Service latency as well as the network contention, resulting in a better service experience for end-users.

Apart from signaling services, a user equipment may also interact with other Network Functions such as third party application functions. By introducing edge computing, these functions are able to be hosted in a decentralized cloud. The main advantage of edge computing lies in the possibility to deploy such decentralized MEC cloud close to the UE's access point of attachment similar to the distributed deployment of some control plane NFs. It is indeed possible for a MEC to be deployed at the RAN edge, in a distributed datacenter or even in a centralized datacenter depending on the service requirement. Edge computing or MEC is not a new technology. To enable the interaction

between MEC system and 5GC control plane, the design approach taken by 3GPP allows the mapping of MEC onto Application Functions (AF) (ETSI, 2018). MEC can thus interact with the 5G system using Network Exposure Function (NEF) that provides information from external application to 3GPP network, or directly with the target 5G NFs if permitted (ETSI, 2020). Edge computing could be a suitable solution for URLLC type scenarios, e.g. collection and analysis of a large amount of information from massive IoT devices such as connected sensors. The application running on a MEC host deployed on the RAN edge could process the data locally and extract the useful information to the central server. Integrated with MEC system, 5G networks need to be resilient to guarantee the Core Network Function availability and to ensure the connection between MEC and 5G Core Network.

2.6. Service function chaining

To deliver an End-to-End service, various Network Functions are required. A service function chain (SFC) defines an ordered set of Network Functions and ordering constraints that must be applied to packets and/or frames and/or flows selected as a result of classification and/or policy to deliver such an End-to-End service. The mechanism of building such function chains and forwarding packets, frames or flows through them is called service function chaining (ITU-T, 2016). From an End-to-End perspective, a SFC defines how a Network Service is implemented. Since the Network Function Virtualisation is applied in 5G networks (see Section 3.1 for more details), the SFC becomes Virtualized Network Function (VNF) chain (see Figure 3 an example of two service function chains). To allocate SFC request on NFV Infrastructure is challenging. The VNF instances should be hosted at the server with enough resources. Some specific rules may define the isolation or co-location of VNF instances. VNFs may have specific behaviors: Some VNFs can be load balancers, thus parallel processing is allowed; Some VNFs may have multiple outputs (next VNF) depending on the attribute of input traversing traffic; Sometimes, the traffic arrived at a VNF cannot be processed immediately, it has to be queued. Taking into consideration all these constraints, the placement of SFC is a complex problem. A lot of works address this problem by proposing different approaches including resolving a shortest path problem (Martini et al., 2015), integer linear programming (ILP) (Baumgartner et al., 2015) or mixed-integer linear programming (MILP) (Dietrich et al., 2017).

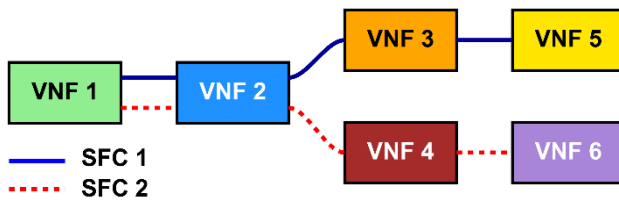


Fig. 3. Two VNF Chains in one network.

Figure 4 shows an example where two service function chains are deployed in the network. Both SFC 1 and 2 start from consuming VNF 1 and VNF 2. Since these two chains serve different services, they consume different VNFs separately thereafter. SFC 1 utilizes VNF 3, 5 and SFC 2 utilizes VNF 4, 6. Each VNF has several instances and they can be deployed on different servers. Multiple SFCs can share the same VNF instance or they consume different instances. In Figure 4, Server 2, 3 are used by both SFCs while Server 1 is used only by SFC 1 and Server 4 is needed only by SFC 2.

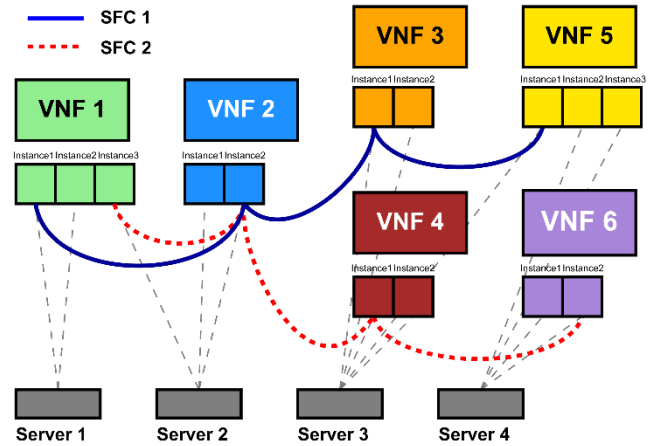


Fig. 4. Deployment of two Service Function Chains.

3. The complexity from the Multilayer perspective

The End-to-End perspective illustrates horizontally the complexity when delivering a Network Service. The multilayer perspective reveals vertically the complexity while orchestrating and managing a Network Service.

3.1. Network function virtualisation

With more and more new services joining the network, the resource allocation and service maintenance become a bottleneck for improving the service performance. In 5G networks, the Network Function Virtualisation (NFV) technology is proposed to solve the problem. This technology reforms the network architecture by separating software from the hardware with the help of virtualisation (Chiosi, 2012). Virtualizing Network Functions enables flexible distribution of hardware resources to improve the service performance and rapid launch of innovative services to generate new revenue sources. It is also an enabler for the formerly mentioned flexible deployment of several 5GC Network Functions, e.g. AMF, and co-locating them with the access network and thus eliminating long-distance data transport (Han et al., 2015). All core NFs and some access NFs are targeted to be virtualized in 5G networks in some scenarios (ENISA, 2020).

With virtualisation, the physical Network Functions become Virtualized Network Functions (VNFs). These VNFs can be deployed on virtual machines (VMs) or containers. The former is a traditional virtualisation environment while the latter is lighter-weight and more

agile. The infrastructure resources including storage, compute and network are virtualized and useable for virtualized layer. Instead of allocating a fixed amount of physical resources, VNFs can allocate dynamically virtual resources according to the service request and traffic in real-time.

More specifically, NFV adopts a three-tier architecture (ETSI, 2014) as shown in Figure 5. At the top is the operation layer, with Business Support Systems (BSS) and Operations Support Systems (OSS) to support various End-to-End telecommunication services. Some processes covered by OSS/BSS include: network management, service delivery, fulfilment, assurance, and billing. Lower down is the Network Service and Network Function layer. Inside this layer, the Virtualized Network Functions are managed by Element Managements (EMs). EM’s role includes security management, fault management for the exposed Network Function services provided by VNFs. At the bottom lies the NFV infrastructure (NFVI). Storage and compute are two main physical hardware resources which are normally pooled. Another physical resource is networking devices including routers and links. The virtualisation layer abstracts the hardware resources and decouples the VNF software from the underlying hardware, ensuring a hardware independent lifecycle for the VNFs. For the majority of current deployments, the virtualisation layer in an NFVI comprises a hypervisor to partition physical servers into VMs and a network controller, typically a Software-Defined Network (SDN) controller, to help partition the physical network that connects the physical servers into multiple virtual networks interconnecting groups of VMs. While the vast majority of NFV deployments is still based on hypervisor technologies, container-based virtualisation (a.k.a. Operating System (OS) virtualisation) is gaining momentum and might become the norm for 5G. Containers provide an isolation capability that allows multiple VNF instances to share the same host OS while virtual machines require a separate guest OS for each VNF instance.

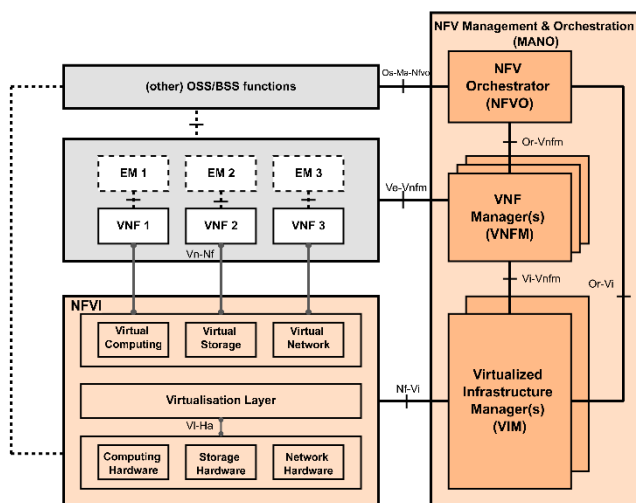


Fig. 5. Network Function Virtualisation architecture.

NFV Management and Network Orchestration (MANO) takes charge of the management of NFVI and orchestrates the allocation of resources needed by the Network Services and VNFs (ETSI, 2016). NFV MANO includes three functional blocks. NFV Orchestrator (NFVO) is responsible in general for the life cycle operations of a Network Service. NFVO functions can be classified into End-to-End resource orchestration, and Network Service orchestration. VNF Manager (VNFM) is in charge of the life cycle operations as well as performance, fault and configuration management of a VNF. Specifically, the managements include instantiation, heal, operation (changing the state), information modification, changing connectivity, scaling and termination. Each VNF manager serves one or multiple VNFs according to the network design. The third block, Virtualized Infrastructure Manager (VIM), involves all life cycle operations of a virtualized resource. Concretely, a VIM controls and manages the interaction of a VNF with physical and virtualized resources including compute, storage and network. Similar to VNFM, multiple VIMs can be deployed in the network.

3.2. VNF deployment

VNF can be VM based or container based depending on the choice of technologies. Traditional VMs virtualizes an underlying computer while the container is lighter, containing the code and dependencies needed, taking less time to image and to run an application.

A VNF is composed of one or multiple VNF components (VNFCs) (ETSI, 2014). A VNFC is a software entity in charge of different functionalities or data bases (in 5G network). Thus, a VNF is mapped to one or several VMs or containers in NFVI servers. It is worth noting that even though there could be multiple VNF components belonging to a same VNF instance, they are not necessarily deployed all on the same host.

3.3. Example of operations in Network Service

Scaling is a typical action to manage a Network Service. Scaling can be categorized into two classes: horizontal scaling and vertical scaling. The former includes scaling in and out, which refers to a process where one or more instances are removed or added. The latter action includes scaling up and down, which refers to a process of adding or releasing resources to or from an existing instance.

Network Service level scaling out and in are important operations during the management of Network Services (3GPP TR32.842, 2015). They may be triggered from OSS, by an operator manually or by some Network Manager-level functions (e.g. Load Balancing) automatically. Then NFVO will receive the request to scale out or in a Network Service instance. For a Network Service, scaling is necessary when one or several of the actual VNF instances in the network is/are overloaded or too redundant for the Network Service. The scaling out in the Network Service level can be done by scaling out or scaling up concerning VNF instances. The service level scaling in can be done by scaling in or scaling down the corresponding VNF instances. Precisely, Network Service can be scaled out (or

in) by expanding (contracting) some of existing VNF instances or by instantiating new (terminating existing) VNF instances. When there is a resource change involved, NFVO will also send a request to VIM to allocate the changed resources.

The VNF level scaling can be triggered from NFVO, VNFEM, EM, OSS, or manually by the operator. This scaling concerns the management of VNF components. The scaling request will be received by VNFEM and it performs the VNFC instantiation or termination procedure for horizontal scaling. In vertical scaling, VNFEM requests an update of the resources for VNFC instances.

In case of threat or failure, an operation can be triggered from different levels of the network. When an overload situation happens on a server, this will cause malfunction on VNFCs relying on it. After failure or overload detection, multiple entities of the network may react competitively as depicted in Figure 6. In case of Kubernetes container, Kubernetes hypervisor may decide an instantiation of one or multiple related Pods (a group of containers) on available servers. VNFEM may scale out or up the concerning VNFC. NFVO may also change the deployment flavor and apply a rule such as instantiating VNF or increasing the capacity allocated to the concerning VNF. The 5G network mechanism at the VNF level may decide to reduce the number of messages sent to the VNF hosted on the overloaded server or redirect the messages. All these decisions are helpful to some extent but not complementary. It is necessary to determine upstream the best entity which must make the decision according to the situation to avoid a bad decision making situation worse. If we find ourselves in a case of a signaling storm, deploying new instances of VNF or VNFC can impact the datacenter at large by saturating again the servers, whereas actions at the 5G network level would have made it possible to resolve the problem.

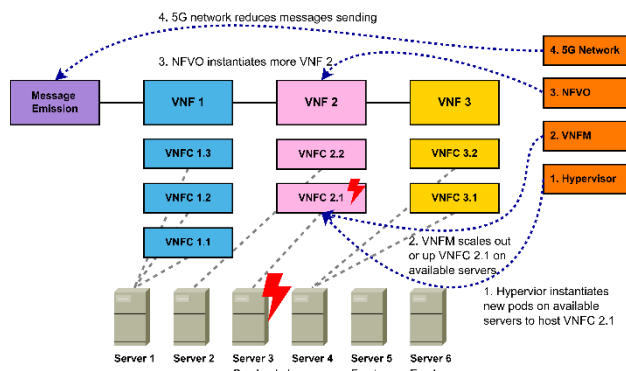


Fig. 6. Four competitive reactions of different network entities to face one overload situation.

4. Procedures of setting up a slice

5G networks enable a variety of Network Services. With the use of network slicing, the application requirements become more challenging and heterogeneous in order to better serve the vertical industry (ITU-T, 2018; Foukas et al., 2015). A use case from vertical industries may be translated into

several Network Services in 5G networks. To set up these Network Services, several procedures such as service selection, resource allocation, are indispensable for Mobile Network Operator (MNO). A detailed example of procedures may be as follows:

- Expression of need and service selection: description of the service and associated requirements referring to an existing catalog for solution
- Communication service request with relevant inputs and performance requirements via Communication Service Management Function (CSMF): this step includes creating service order (Sending Customer Facing Service information to Service Order Management), updating resource order (sending Resource Facing Service information to Resource Order Management) and sending a request to Network Slice Management Function (NSMF) in order to create a Network Slice Instance (NSI) for this Service Profile
- Slice selection via NSMF: NSMF determines network Slice subnet requirements (slice profile for Subnet includes RAN, Core Network and Transport Network subnets) and the associated Network Slice Subnet Management Function (NSSMF); Secondly, NSMF sends a Network Slice Subnet Instance (NSSI) allocation request to NSSMF
- Subnet creation via NSSMF: a NSSMF takes charge of checking the feasibility of Network Slice Subnet Requirements, creation of a new NSSI as well as a New RAN Network Slice with RAN Network Functions (in RU, DU, CU), determination of Network Service Descriptor for Core Network NSSI components
- Service orchestration via NFV-MANO: NFVO derives a SFC, the location and behaviors of the VNFs from the requirements; NFVO interacts with the VNFEM of each VNF and with the VIM of each datacenter; Each VNFEM instantiates VNF by taking into account redundancy requirements, affinity or anti-affinity rules to select the best server to host the VNF components; VNFEM sends the instantiation request to the VIM; VIM creates containers for VNF components, allocates resources for each container and ensures the connectivity between the servers inside the datacenter
- Transport Network Subnet creation via NSSMF: NSSMF derives requirements for Transport Network NSSI component
- Service orchestration via NFV-MANO in Transport Network: NFVO derives Network Slices Virtual Links Descriptor and interacts with WAN Infrastructure Manager to connect RAN and Core Network Point of Presence. The transport network will ensure the connectivity of all required functions from RAN or Core Networks.

5. Resilience challenges in 5G network

Resilience is a relatively new field in system engineering that drew great attention over the last decade. Resilience could be defined as the ability of a system to prepare and

plan for, absorb, recover from, and more successfully adapt to adverse events. (National Academies Committee on Increasing National Resilience to Hazards and Disasters, 2012).

From ENISA Telecom Services Security Incidents 2019 Annual Analysis Report, system failures, human errors and natural phenomena are the three main causes of telecommunication incidents (ENISA, 2020). More than half of the telecom security incidents were caused by system failures. The most common system failures are hardware failures and software bugs. This type of adverse events often happens on one single element, however other elements of the communication network may also be victims as a result of failure propagation due to interdependencies in telecommunication networks (Martins et al., 2017). Human errors normally result from an imperfect system design or a wrong configuration. For natural phenomena, the main characteristic is the broad effect in scope. For example, a blizzard would possibly result in multiple network nodes and links failures in the affected region. Other kinds of failures including malicious attacks are rare compared with the three main failures, thus are not the focus in our approach of resilience challenges in 5G network.

A resilient 5G network should be able to offer services with high Quality of Service (QoS) at all time regardless of the adverse events. QoS is the ability of a service to comply with quality requirements and service level as agreed (or targeted) with the end user. The QoS is often interpreted into important performance parameters of the telecommunication system, typically referred to Key Performance Indicators (KPIs) (Kukliński et al., 2019). As 5G is supposed to be a vertical service delivery model, the challenge of resilience in 5G systems resides in the fact that KPIs and service requirements vary from one case to another. These KPIs may include latency, availability, throughput, etc. In the use case of tele-action in power systems, the End-to-End service latency is the main KPI, which should be less than 50 milliseconds in order to guarantee the functioning of the grid network (ENEDIS, 2020). In another use case of autonomous vehicles in manufacturing environments, network availability should be higher than 99.90% and network End-to-End latency should be lower than 10 milliseconds (5G EVE, 2018).

To evaluate the performance and resilience of 5G Network Services, recent works focus on two types of situations, namely in the design process and system recovery process. With the NFV environment, Service Function Chain (SFC) becomes the carrier of Network Services. Then the problem is to evaluate and optimize the SFC deployment. In the SFC design process, it aims to prevent failure before it happens by analyzing for example necessary redundancy on each element of the chain and verify the KPIs on the chain as well as the redundant ones. In the recovery process, it is important to analyze the cost and time to provision the backup elements in the chain. Given the complexity of 5G, the problem needs to be solved by taking into consideration of both End-to-End and multi-layer perspectives. From the End-to-End perspective, a SFC should include the Radio Access Network, Core Network,

Transport Network, and the Edge Computing and Data Network if applicable. While modeling the network, we should be capable to answer how is RAN deployed and where are RU and BBU deployed; what are the Control Plane Core Network Functions engaged in this SFC; what technologies are used on Fronthaul, Midhaul, Backhaul and Backbone; what are the elements co-located in the network. On the other hand, the multi-layer perspective emphasizes the complexity of SFC management. The operations such as network element scaling and failure recovery should be modeled. The competitive actions from different layers as presented in Section 3.3 will also be a challenge for resilience analysis.

The mainstream researches optimize the SFC deployment while subjecting to the resources, placement and performance constraints. These optimizations are mostly based on ILP or MILP models. The optimization goals vary from one to another, such as minimizing the bandwidth usage across the network (Qu et al., 2017), minimizing the total cost of deploying all network slice requests (Da Silva Coelho et al., 2020), minimizing the cost of protecting the SC against single failure (Carlinet et al., 2020), jointly minimizing the overall deployment cost and service delay (Leivadeas et al., 2019). These works simplify the Transport Network and Core Network and neglect the complex RAN. Only a few of them consider the system recovery process and these models often neglect the interdependency between each layer, the existence of MANO, the interaction between NFV-MANO and VNFs. For example, in the recovery process, the failure on MANO actually will block the VNF level or service level scaling.

Some recent works based on Petri Net have studied the recovery process (Rui et al., 2020), the decomposition of VNF (Di Mauro et al., 2017), Network Function behaviors (Schneider et al., 2019), NFV-MANO structure (Tola, et al., 2019) in the telecommunication network. The models based on Petri Net and its extensions seems to be a promising tool in Network Service evaluation since it can better describe the complex 5G network. Petri Nets' main attraction as a modeling formalism is how the basic aspects of concurrent systems are identified both conceptually and mathematically (Bonet, 2007). The marking of the state of a Petri Net model shows the state of the telecommunication system. The transition of a Petri Net model represents an action, e.g. scaling, failure and recovery. Its extensions such as Colored Petri Nets, Timed Petri Nets and Stochastic Petri Net enrich the capacity of the model and make it possible to measure the performance of the Network Service, e.g. availability and latency.

6. Conclusion

In this paper, we have introduced the complexity in 5G networks from both End-to-End and multi-layer perspectives. Some use cases are given to further explain the complexity in setting up a Network Service. By implementing new technologies, in particular NFV, 5G networks are becoming more flexible but also more complex. Therefore, more and more resilience challenges

are awaiting to be taken before introducing 5G networks into new scenarios.

As we noted, modeling the network and its complexity is an important step for evaluating Network Service performance. We propose Petri-Net as a promising tool for the 5G network performance analysis. The work on modelling 5G networks to evaluate the resilience of an End-to-End service related to verticals, is in progress.

References

- Specific Requirements. Part 11: Wireless LAN Medium Access Control, Standard IEEE 802.11ad, 2012.
- Patriciello, N., S. Lagén., B. Bojović and L. Giupponi (2020). NR-U and IEEE 802.11 Technologies Coexistence in Unlicensed mmWave Spectrum: Models and Evaluation. *IEEE access* 8, 71254-71271.
- 3GPP (2021 (accessed April 2021)). *TS 23.501 V17.0.0 System architecture for the 5G System (5GS)*.
- Mademann, F. (2018). The 5G system architecture. *Journal of ICT Standardization* 6 (3), 77-86.
- 3GPP (2021 (accessed April 2021)). *TS 28.530 V16.9.0 Management and orchestration; Concepts, use cases and requirements*
- 5G-SMART (2020 (accessed April 2021)). 5G Common Terminology. *5G Smart Manufacturing*. <https://5gsmart.eu/wp-content/uploads/5G-SMART-common-terminology.pdf>
- ETSI (2019 (accessed April 2021)). *GS MEC 001 V2.1.1. Multi-access Edge Computing (MEC) Terminology*.
- Banchs, A., D. M. Gutierrez-Estevéz, M. Fuentes, M. Boldi, and S. Proveddi (2019). A 5G mobile network architecture to support vertical industries. *IEEE Communications Magazine* 57(12), 38-44.
- ITU-R (2015). Framework and overall objectives of the future development of IMT for 2020 and beyond. *Radiocommunication Sector of ITU*.
- Wind River (2017(accessed April 2021)). vRAN: The Next Step in Network Transformation.
- ETSI (2018 (accessed April 2021)). MEC in 5G networks. *ETSI White Paper No. 28*.
- ETSI (2020 (accessed April 2021)). *GR MEC 031 V2.1.1. Multi-access Edge Computing (MEC): MEC 5G Integration*.
- ITU-T (2016). Deployment models of service function chaining. *ITU Y-series Recommendations—Supplement 41*.
- Martini, B., F. Paganelli, P. Cappanera, S. Turchi, and P. Castoldi (2015). Latency-aware composition of virtual functions in 5G. *Proceedings of the 2015 1st IEEE Conference on Network Softwarization (NetSoft)*, 1-6.
- Baumgartner, A., V. S. Reddy, and T. Bauschert (2015). Mobile core network virtualization: A model for combined virtual core network function placement and topology optimization. *Proceedings of the 2015 1st IEEE conference on Network Softwarization (NetSoft)*, 1-9.
- Dietrich, D., C. Papagianni, P. Papadimitriou, and J.S. Baras (2017). Network function placement on virtualized cellular cores. *2017 9th International Conference on Communication Systems and Networks (COMSNETS)*, 259-266.
- Chiosi, M. (2012). Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges & Call for Action. *ETSI White Paper*.
- Han, B., V. Gopalakrishnan, L. Ji, and S. Lee (2015). Network function virtualization: Challenges and opportunities for innovations. *IEEE Communications Magazine* 53(2), 90-97.
- The European Union Agency for Cybersecurity (ENISA) (2020). *ENISA threat landscape for 5G networks – Updated threat assessment for the fifth generation of mobile telecommunications networks (5G)*
- ETSI (2014 (accessed April 2021)). *GS NFV-MAN 001 V1.1.1 Network Function Virtualisation (NFV); Management and Orchestration*.
- ETSI (2014 (accessed April 2021)). *GS NFV-SWA 001 V1.1.1 (2014-12) Network Functions Virtualisation (NFV); Virtual Network Functions Architecture*
- 3GPP (2015 (accessed April 2021)). *TR 32.842 V13.1.0. Telecommunication management; Study on network management of virtualized networks*
- ITU-T (2018). Service function chaining in mobile networks. *Telecommunication standardization sector of ITU*.
- Foukas, X., G. Patounas, A. Elmokashfi, and M.K. Marina (2017). Network slicing in 5G: Survey and challenges. *IEEE Communications Magazine* 55(5), 94-100.
- Committee on Increasing National Resilience to Hazards and Disasters (2012), *Disaster Resilience: A National Imperative*, Natl. Acad. Press.
- The European Union Agency for Cybersecurity (ENISA) (2020). *Telecom Services Security Incidents 2019 Annual Analysis Report*.
- Martins, L., R. Girao-Silva, L. Jorge, A. Gomes, F. Musumeci, and J. Rak (2017). Interdependence between power grids and communication networks: A resilience perspective. *DRCN 2017-Design of Reliable Communication Networks; 13th International Conference*, 1-9.
- Kukliński, S., and L. Tomaszewski, (2019). Key Performance Indicators for 5G network slicing. *2019 IEEE Conference on Network Softwarization (NetSoft)*, 464-471.
- ENEDIS (2020). *Description et étude des protections de découplage pour le raccordement des Installations de Production raccordées au Réseau Public de Distribution*. https://www.enedis.fr/sites/default/files/Enedis-NOI-RES_13E.pdf
- 5G EVE (2018). *Deliverable D1.1 Requirements Definition & Analysis from Participant Vertical Industries 5G European Validation platform for Extensive trials*. <https://www.5g-eve.eu/wp-content/uploads/2018/11/5g-eve-d1.1-requirement-definition-analysis-from-participant-verticals.pdf>
- Qu, L., C. Assi, K. Shaban, and M. J. Khabbaz (2017). A reliability-aware network service chain provisioning with delay guarantees in NFV-enabled enterprise datacenter networks. *IEEE Transactions on Network and Service Management* 14(3), 554-568.
- Da Silva Coelho, W., A. Benhamiche, N. Perrot, and S. Secci (2020). On the impact of novel function mappings, sharing policies, and split settings in network slice design. *2020 16th International Conference on Network and Service Management (CNSM)*, 1-9.
- Carlinet, Y., N. Perrot, and A. Alves-Tzitas (2019). Minimum-Cost Virtual Network Function Resilience. *INOC 2019*.
- Leivadreas, A., G. Kesidis, M. Ibnkahla, and I. Lambadaris (2019). Vnf placement optimization at the edge and cloud. *Future Internet* 11(3), 69.
- Bonet, P., C. M. Lladó, R. Puijaner, and W. J. Knottenbelt (2007). PIPE v2. 5: A Petri net tool for performance modelling. *23rd Latin American Conference on Informatics CLEI*, 50-62.
- Rui, L., X. Chen, Z. Gao, W. Li, X. Qiu, and L. Meng (2020). Petri Net-Based Reliability Assessment and Migration Optimization Strategy of SFC. *IEEE Transactions on Network and Service Management* 18(1), 167-181.