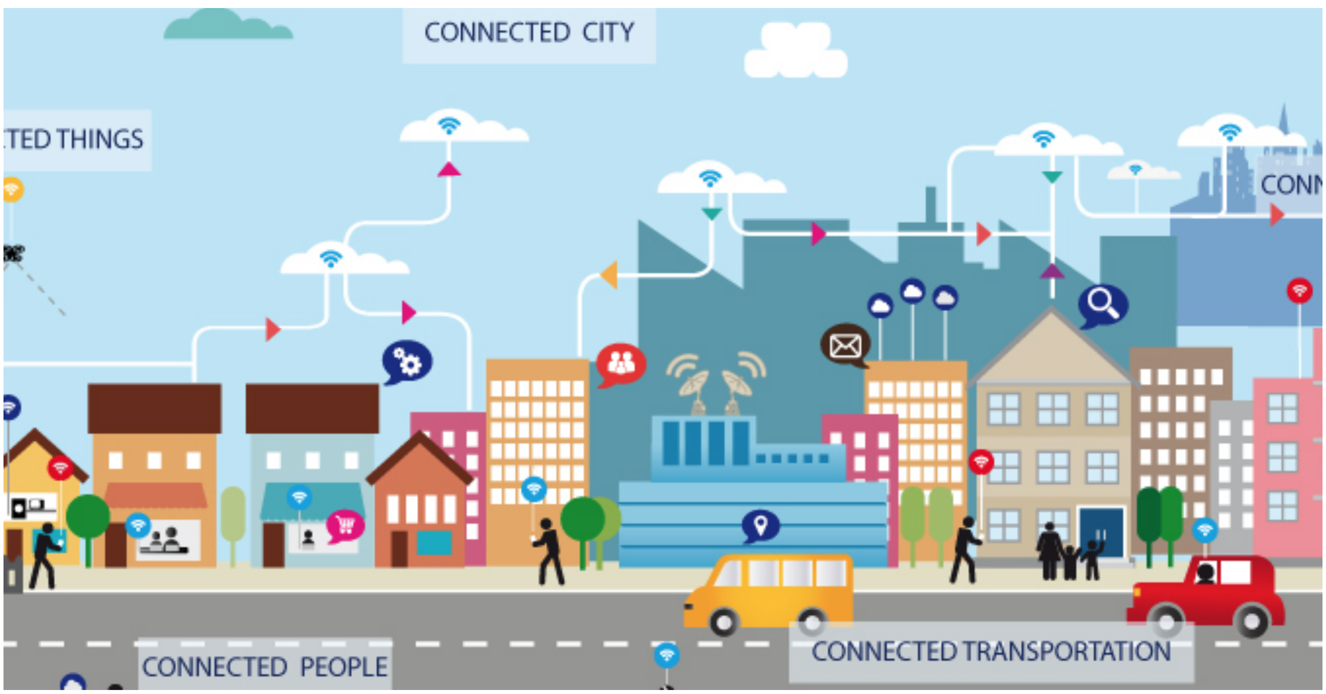
5G NETWORK SLICING WHITEPAPER

FCC Technological Advisory Council

5G IoT Working Group



Kevin Sparks, Marvin Sirbu, Jack Nasielski, Lynn Merrill, Kevin Leddy, Padma Krishnaswamy, Walter Johnston, Russ Gyurek, Brian Daly, Mark Bayliss, John Barnhill, Kumar Balachandran

TABLE OF CONTENTS

Executive Summary 1
Overview of Network evolution from service specific networks to converged networks 3
Overview of what slicing means related to 5G – The Fundamentals 9
Network slicing and 3GPP Activities 14
Technical deep dive on slicing 17
Services related to slicing 24
Summary 25
Slicing Whitepapers to Reference: 26
Glossary - Definitions 27

FIGURES

Figure 1: Evolution of mobile digital telecommunication systems 8
Figure 2: 3GPP RAN Progress on "5G" March 2016 10
Figure 3: Service-based network slicing. 11
Figure 4: Principles of slicing. 12
Figure 5: RAN Chairman's Summary from RAN#79, March 2018 14
Figure 6 17
Figure 7: Lifecycle phases of an NSI 18
Figure 8: 19
Figure 9 21
Figure 10 22
Figure 11 24

Executive Summary

5G will bring many benefits to both residential and enterprise customers. The technologies tied to 5G are transformational in terms of the impact to the communications network: massive bandwidth, sliceability, ultra-low latency, massive scale IoT, and imbedded security. A network slice is defined as *a logical (virtual) network customized to serve a defined business purpose or customer, consisting of an end-to-end composition of all the varied network resources required to satisfy the specific performance and economic needs of that particular service class or customer application.* This paper focuses on 5G network slicing.

Slicing is not a new concept. Virtual network capabilities have been part of packet networking for decades. However, 5G deployments will extend this virtualization to an end-end/top-to-bottom functional scope and imbed slicing as a core functioning part of the network. The benefits include internal service provider network management uses; the ability to differentiate broad classes of services that require certain characteristics or resource parameters; providing a virtual service provider network across another physical network operator; providing customers the ability to customize a virtual network to support their operations; traffic splitting across 5G, 4G and Wi-Fi networks; etc. Operators will utilize slicing to optimize network management from core to customer.

In addition, the benefits of slicing for specific verticals are quickly emerging. Ultra- reliable low latency communications (URLLC) is focused on providing support for critical infrastructure and applications such as connected transportation and real-time processes such as in manufacturing. IoT has numerous verticals and each of those verticals have varied network requirements. Slicing provides the ability to support these vertical use cases by creating a tailored virtual slice of the network from end to end.

Standards organizations such as 3GPP are actively leading the effort to define slicing capabilities at the technical and architectural levels. This includes work on the management and orchestration of network slicing as it applies to next generation networking. Slicing specifications will have a major update by 3GPP in release 16 at the end of CY 2019.

The paper is intended to provide a background on what slicing provides, its origins and evolution, how it will be implemented, and some of the many use cases. As slicing is an evolving technology, it is recommended that readers continue to monitor the standards progress as well as service deployments. The FCC TAC team is optimistic on the value slicing brings to communication networks.

Recommendation*: The FCC 5G/IoT working group has made a formal recommendation for the FCC to take action on providing "regulatory clarity" as it relates to slicing services. (assuming approval, paper will be published at fcc.gov/tac-reports-and-papers) Lack of clarity may hamper investments and or limit innovation related to new services.*

Overview of Network evolution from service specific networks to converged networks

Communications networks are intrinsically dominated by the underlying economics of the technologies used to construct such networks. Prior to the emergence of broadband transport and packet switching in the 1980s, economics dictated that each network be optimized to support the primary type of service it would support. Prior to widespread adoption of the public Internet, voice communication service was, by far, the dominant service type. Transport networks were optimized to support the requirements of the Public Switched Telephone Network (PSTN), first over analog then digital channelized voice trunks. In addition, other specialized networks were developed for the transport of other types of services though such networks lacked the scope and scale of the ubiquitous voice network. For example, local telephone companies had developed intracity video networks for the transport of broadcast video traffic. A nationwide digital service network (Digital Data Service) was launched by AT&T among major cities in 1974 which would eventually offer switched data rates up to 56 kbps. Later dedicated packet networks using Frame Relay and ATM were built to support higher capacity enterprise and network data needs. However, until the large-scale introduction of digital broadband services and fiber optic transport systems in the 1980s, the dominant long-distance transport mechanism was based on analog technology, whether wireline (copper, coax), wireless or satellite, and the dominant service class remained voice traffic. Meanwhile, a completely separate network for the delivery of video traffic emerged based on analog transmission over local coaxial cable networks, with content distributed by satellite to municipal headends.

From a network engineering perspective, these purpose-built networks were justifiable in their day as engineers understood both the advantages and disadvantages of networks optimized to transport a single type of service, be it voice, video or data. While this approach produced, as a prime example, an efficient and high-quality voice network, it severely constrained the ability of that network to transport other service types. For example, data could be transported across switched voice networks but only using data modems which translated the data stream into an analog signal constrained to a 3.5 Khz bandwidth, matching the attributes of a voice signal for which the network was designed, which severely limited maximum data transfer rates.

Since the choices were initially limited to circuit-switched technology, as engineers evaluated designs which could offer more service flexibility, the economics dictated that the network be optimized to support voice service either in its analog form or its digital equivalent of a 64 kb/s channel.

The introduction of the Arpanet—a packet network funded by DARPA and first deployed in 1969—heralded a significant new phase of research on multi-service or converged networks. Arpanet and similar research efforts of that era such as Cyclades, Tymnet, X.25 and others[1] marked the emergence of *packet switching*.

With packet networks, digital data is carried in a flexible format across the network as compared to earlier TDM technology. Depending upon coding, a data packet could represent any form of encodable information: text, image, video or voice. Services would largely be provided by intelligent endpoints external to the communications network. Early experiments in voice transmission across Arpanet began in 1972 and by the time of the introduction of optical transmission systems in the 1980s, there was an existing body of engineering knowledge regarding using packet networks as multi-service networks.

The introduction of optical high-speed digital technology beginning in 1980 with deployment of the first commercial optical system within the United States, changed the underlying economics of communication networks. For the first time, a wide-ranging digital backbone network provided both abundant bandwidth and the most economic choice for all forms of service transport. For example, the significant economies in transmission cost reaped through optical broadband transport meant that even video services could be transported economically across packet networks at rates affordable to the consumer.

The availability of a very high speed, converged backbone led to accompanying advances in high speed access to the consumer (residential or business). With this advent of broadband access services to the consumer in the late 1990s, a

[1] Roberts, Lawrence E., "The Evolution of Packet Switching," *Proceedings of the IEEE*, 66, 11, Nov 1978, pp 1307-1313.

converged, high speed wireline network foundation— for the first time—was in place and capable of efficiently supporting a wide range of services on a single platform.

As network physical layer technologies in the backbone network evolved along the lines described, the ARPANET in parallel had transitioned from primarily defense and academic use, to an "Internet" and a variety of associated consumer offerings.[2]

The progressive diversification and commercialization of the Internet in the 1990s also led to it enabling an increasingly wider array of services: email, chat, file transfer, streaming audio and video services among others.

Realizing the potential of the optical transport on the backbone required an associated upgrade of the access network. Physical layer technologies such as Digital Subscriber Line (DSL) over telephone wires, fiber-to-the-home (FTTH), the evolution of cable to two-way hybrid fiber-coax (HFC), local area networks (LANs) and WiFi brought higher speeds directly to end user devices. With broadband access services offered to consumers in the late 1990s onward, for the first time a network foundation was in place capable of efficiently supporting a wide range of services on a single platform. Over the last two decades, network performance, driven by Moore's law, has increased apace with *fixed network* consumer access rates now exceeding 100 Mbps. In parallel, mobile networks have evolved from providing only analog voice to providing tens or even hundreds of Mbps over packetized platforms.

The network specialization intrinsic to voice and video networks had formed a business barrier between telephone network operators and cable system operators. However, it had been apparent for some time that digital broadband would remove historical barriers to entry. In addition, the open architecture of the Internet which was the driving force for digital broadband services to the consumer, also facilitated a new class of service provider — "Over the Top" (OTT) operators. It permitted OTT companies to offer services to consumers and

[2]Zimmerman, K.A. and Emspak, J., "Internet History Timeline from Arpanet to the WWW, " *LiveScience*, June 27, 2017. Available at: https://www.livescience.com/20727-internet-history.html

businesses over the Internet, supported by broadband internet access service (BIAS) providers. Essentially, for OTT businesses, the Internet provided access to a potential market consisting of all broadband Internet subscribers

Two independent development cycles had converged to produce a profound change in the communications infrastructure of the United States and the world.

Technology evolution targeted at driving down the cost of voice networks ultimately resulted in both a fixed and mobile broadband network infrastructure capable of delivering nearly any service imaginable.

Research work funded by the government focused on developing resilient networks delivered ARPANET, a research network that demonstrated the ability of packet networks to support a wide range of services and led, via its successors such as the NSFnet in the United States, and similar academic networks elsewhere, to the commercial Internet. Broadband consumer access to the Internet, beginning in the late '90s, was the final link in the chain as it permitted tying business and consumers to both a fixed and mobile broadband converged network infrastructure that continues to catalyze new services and innovation in our economy.

In 1980 AT&T introduced the idea of telephone switches (Service Switching Points) which would route calls in response to instructions determined by software running on centralized processors described as Network Control Points (NCPs)[3], and later as Service Control Points (SCPs).[4] This abstraction was the first instance of a software defined network which would lead, in the case of the telephone network, to the ability to create virtual private voice networks for individual firms, that ran over the shared public network infrastructure, and referred to at the time as "Software Defined Networks."

But the early wireline Internet, like the early telephone network, used distributed intelligence. In the early 2000's the NSF supported the creation of Planet Lab, a network infrastructure that could be sliced into multiple virtual networks, though

[3] John Lawser, "Common Channel Signaling Network Evolution," *AT&T Technical Journal*, **66**, 3, 1987 pp. 13-20.
[4] Richard Robrock, "The Intelligent Network—Changing the Face of Telecommunications," *Proceedings of the IEEE,* **79**, 1, Jan 1991, pp. 7-19

routing was still decentralized.[5] In 2005, Greenberg, et al proposed a rearchitecting of Internet control that would separate routing decision-making from the packet forwarders that would implement routing decisions.[6] In 2008 McKeown and others introduced OpenFlow, a standard for communication between routers and a centralized network controller.[7] The Software Defined Networking (SDN) concept applied to packet networking and the OpenFlow concept led to rapid development of networking solutions which separated forwarding devices from routing control logic.[8] Programmable routing control led to the creation of myriad software defined virtual networks running on commodity packet forwarding hardware in data centers or as overlays over the Internet to enterprise locations. In this decade, commercial ISPs are using these principles to offer virtualized packet networks to corporate customers on demand over their wired networks in an echo of the voice SDNs of the 1990s and this trend is likely to continue as these providers move to incorporating new cloud-based network functions. For commercial mobile BIAS providers, emerging standards for 5G wireless networks are standardizing the capability of creating virtualized slices of the wireless radio access network and the packet core (Slicing is defined more precisely in the following section). Converged 5G cores will be able to terminate traffic from a variety of wireless and wireline access technologies. Because of this convergence, end users will be able to receive reliable and consistent service that is network agnostic.

The separation of the control plane functions from the user (traffic) plane forwarding functions has contributed to a parallel and complementary concept of 'software-ization' of functions, impacting areas well beyond the routing domain. This concept has been applied to all types of network devices from CPE boxes and

---

[5] Brent Chun, David Culler, Timothy Roscoe, Andy Bavier, Larry Peterson, Mike Wawrzoniak, and Mic Bowman, "PlanetLab: an overlay testbed for broad-coverage services," *SIGCOMM Comput. Commun. Rev.* **33**, 3 (July 2003), 3-12

[6] Albert Greenberg, Gisli Hjalmtysson, David A. Maltz, Andy Myers, Jennifer Rexford, Geoffrey Xie, Hong Yan, Jibin Zhan, and Hui Zhang, "A clean slate 4D approach to network control and management." *SIGCOMM Comput. Commun. Rev*. **35**, 5 (October 2005), 41-54.

[7] Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner, "OpenFlow: enabling innovation in campus networks," *SIGCOMM Comput. Commun. Rev*. **38**, 2 (March 2008), 69-74.

[8] Nick Feamster, Jennifer Rexford, and Ellen Zegura. 2014. "The road to SDN: an intellectual history of programmable networks," *SIGCOMM Comput. Commun. Rev*. **44**, 2 (April 2014), 87-98.

firewall gateways to packet core, IMS service complexes, and even RAN baseband processing. The result is that dedicated 'appliance' devices that are difficult to upgrade have been replaced by either general hardware devices whose actions can be rapidly changed by easily upgradeable centralized control software, or by fully implementing the function in software running on general purpose compute servers. This shift toward software running on commodity compute servers is referred to as Network Function Virtualization (NFV).

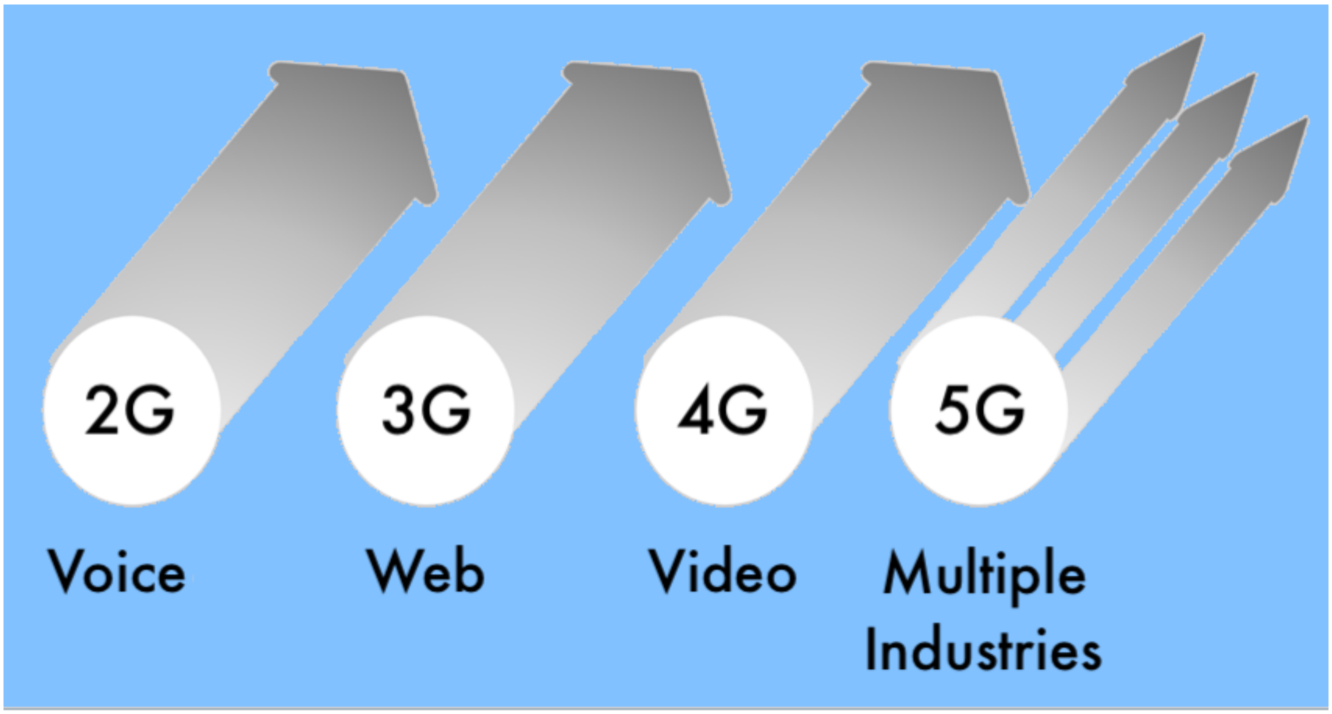


Figure 1: The evolution of mobile digital telecommunication systems has evolved within 5G to support multiple industry needs, in the public and private sphere, creating a need for separation of concerns between various use cases; network slicing provides the semantics for realizing those concerns.

The development of mobile technologies has mirrored the development of the PSTN, transitioning from voice networking to Internet technologies happening over three decades as opposed to the 18-decade history of the telecom industry. Figure 1 is a view of the foci of the various generations of mobile technologies spanning the modern telecommunications age.

With 5G, manufacturers must acquire the ability to provide service creation and delivery at scale across multiple industries. Network slicing is the paradigm that will allow them to achieve that objective.

Overview of what slicing means related to 5G – The Fundamentals

This section provides a general introduction to definitions and concepts associated with slicing and introduces the related 5G standards that have been adopted or are in development.

A Network Slice is a logical (virtual) network customized to serve a defined business purpose or customer, consisting of an end-to-end composition of all the varied network resources required to satisfy the specific performance and economic needs of that particular service class or customer application. The ideas in play in developing and progressing the 'slice' concept draw on a progression of similar but simpler parallels in preceding network architectures including IP/Ethernet networking services (VLANs, IP VPNs, VPLS, etc.), and broaden the scope to include a wide range of access and core network functions from end-to-end and from the top to the bottom of the networking stack.

Network slicing offers a conceptual way of viewing and realizing service provider networks by building logical networks on top of a common and shared infrastructure layer. Network slices are created, changed and removed by management and orchestration functions, which must be considerably enhanced to support this level of multi-domain end-to-end virtualization.

Network slicing is a foundational concept for '5G era' wireless and wireline networks and can be utilized in many different ways. It is useful to differentiate between a few salient levels of usage of network slicing, which is also likely to map to phases of adoption:

- Network Slicing can be used for operational purposes by a single network operator, to differentiate characteristics and resources for different broad classes of services
- Network slicing can be used by a service provider seeking to establish a virtual service provider network over the infrastructure of a physical network operator
- Network slicing can allow individual end customers (enterprises) to be able to customize a virtual network for their operations and consume these network

resources in a more dynamic way similar to today's cloud services (i.e. dynamically varying scale, or for temporary needs).

- Network slicing can allow for "traffic splitting" across networks (5G, 4G, and WiFi via hybrid fiber-coax).

3GPP is the principal standards organization engaged in architecture development for 5G. Several iterations of standards releases have established a foundation for the current phase of slice-specific activity.

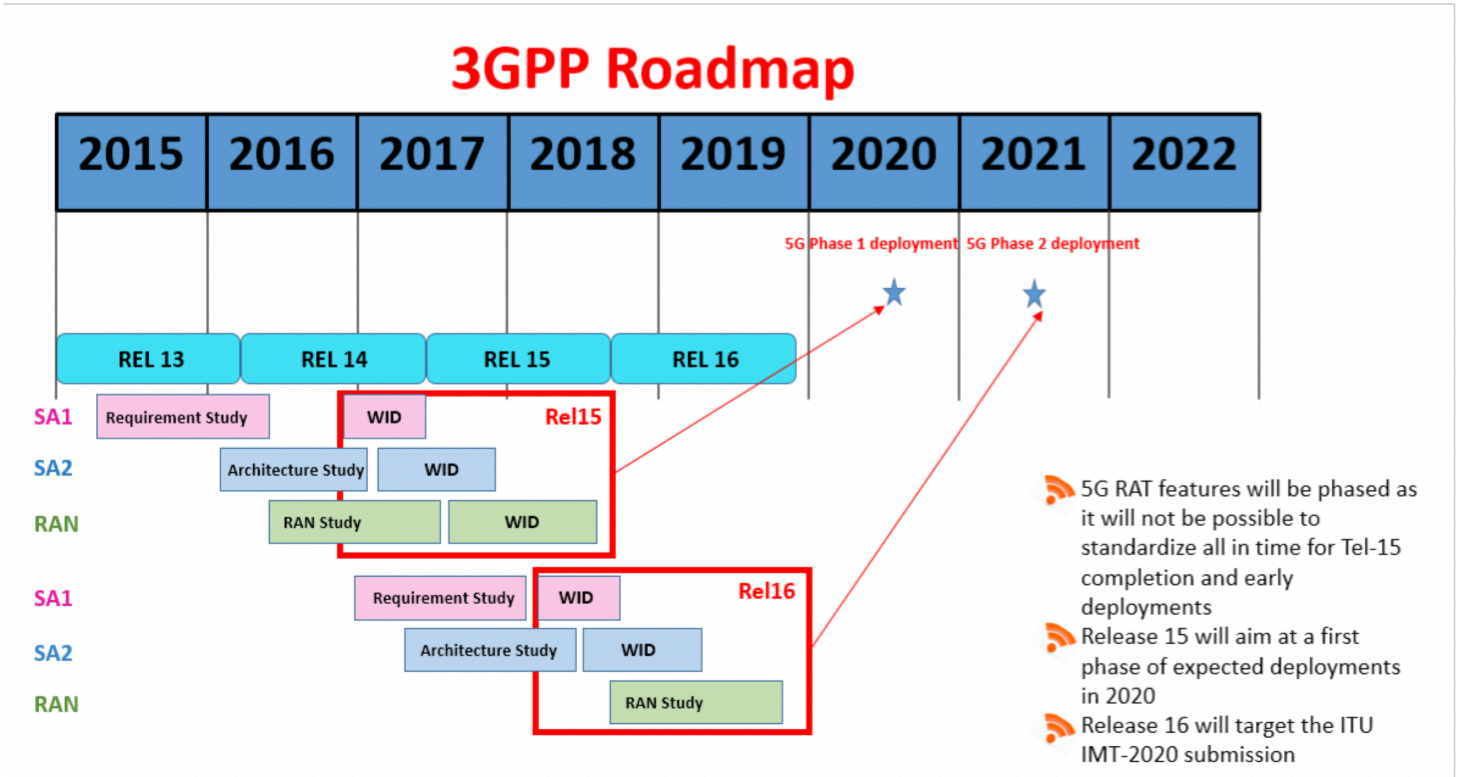


Figure 2: 3GPP RAN Progress on "5G" March 2016

Since 3GPP R13 Dedicated Core Network (DÉCOR), 3GPP has been defining methods which enable different types of devices to be served by different packet core sub-networks instances. The path to network slicing functionality has been laid by enhanced DÉCOR (eDÉCOR) in R14 standards and fully realized with the work on network slicing within the R15 System Architecture for the 5G System (3GPP TS 23.501).

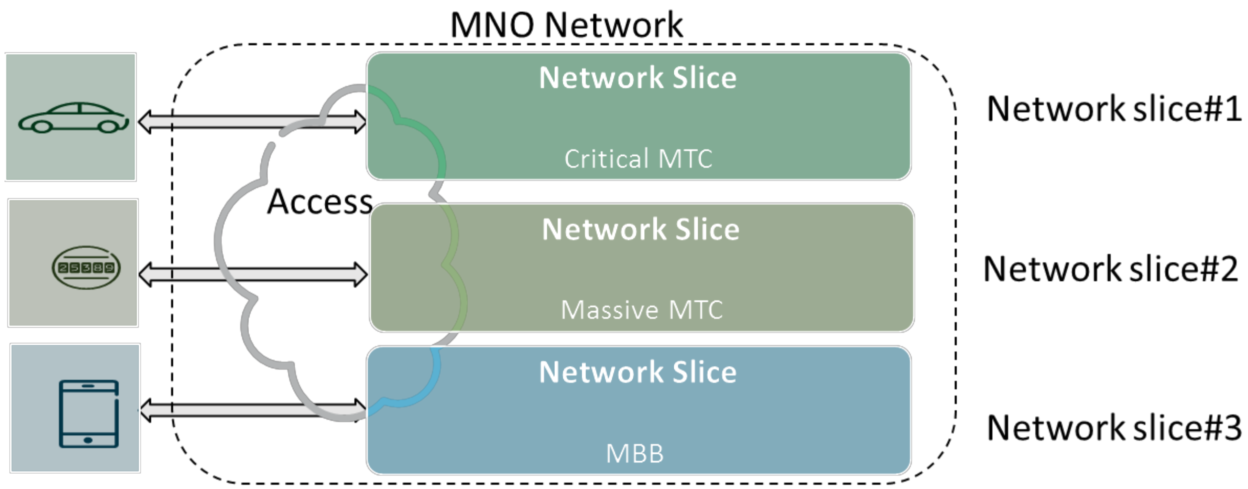


Figure 3: Service-based network slicing in the Release 15 3GPP specifications. Source Ericsson

The concept of a network slice as defined in 5G adds the options to include IMS-, RAN-, UDM and transport resources. In relation especially to the first level above (see Figure 2), the 3GPP has conceived three service-based slices based on the various 5G usage scenarios[9].

These are the Slice/Service types (SST) currently specified by 3GPP. Further customization/specialization is likely for more specific usage scenarios. Note that the GSMA has been working to define Network Slicing Templates (NEST) to define slices that will support specific applications and vertical industries [ref: GSMA papers 3, 4.] GSMA has done primary market research studies talking directly to potential network slice customers. They are currently working to define how slicing templates can be used to parameterize a network slice and allow for easier mobility and roaming across operators.

---

[9] ref: 5G Americas Rel. 14-16 report; *"URLLC and network slicing in 5G enterprise small cell networks";* Issue date: 19 February 2018; Version: 199.10.01

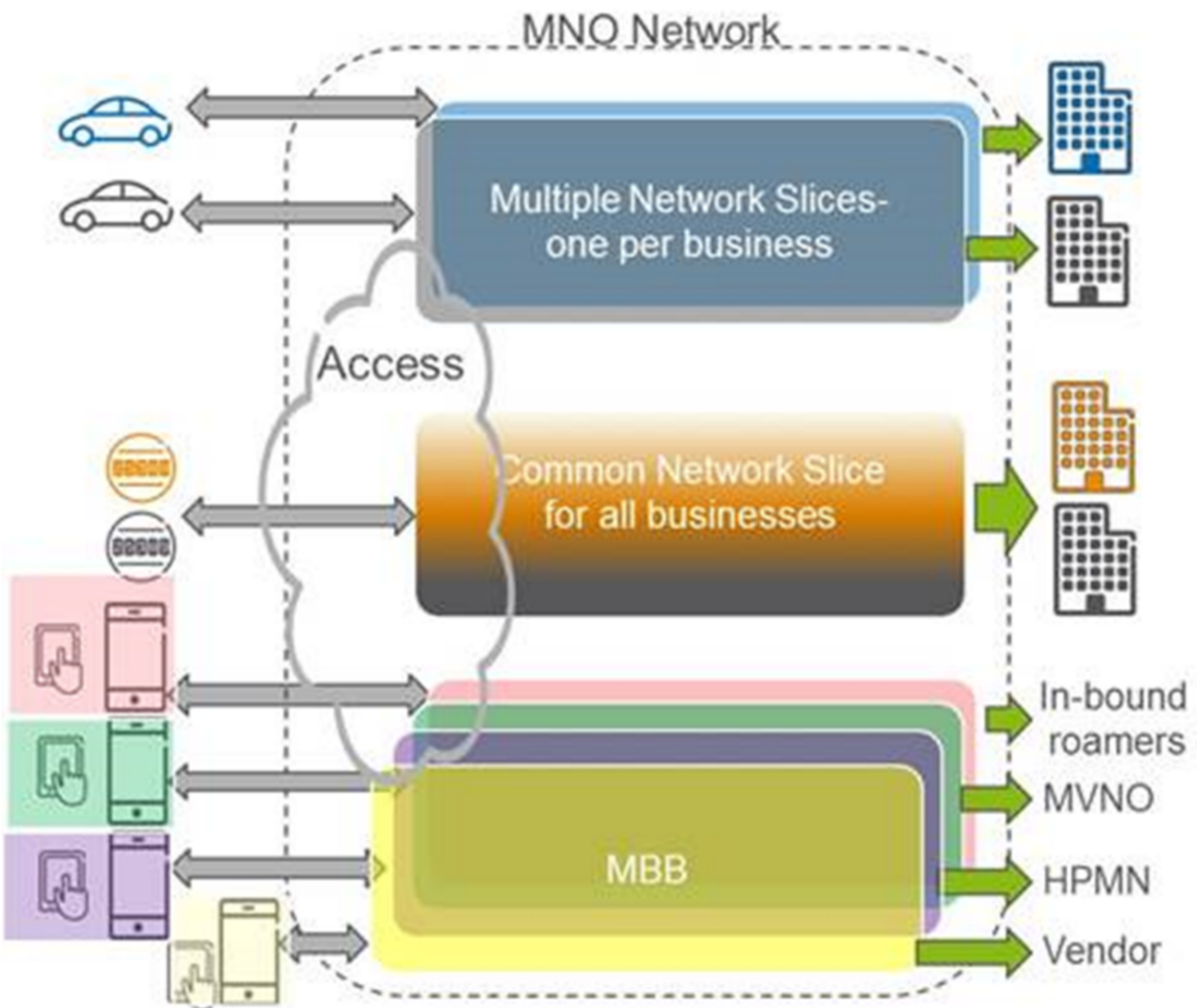


Figure 4: The principles of slicing can be used to realize different objectives can (a) separate business needs, (b) isolate use cases such as Massive IoT, or (c) enable a separation of service requirements within an operator's core network and RAN. Source Ericsson

Network Slices consist of resources composed into an end-to-end service delivery construct. This could be physical resources, either a share or profile allocated to a slice or even dedicated physical resources if motivated. Slices also consist of logical entities such as configured network functions, management functions, VPNs etc. Resources (physical or logical) can be dedicated to a slice, i.e. separate instances, or they could be shared across multiple slices. These resources are not necessarily all produced within the provider, some may in fact be services consumed from other providers, facilitating e.g. aggregation, cloud infrastructure, roaming etc.

In its initial incarnation, network slicing is more likely to be instantiated in a manner similar to the provisioning of VLAN or VPN services for business customers. Over time, network slicing will allow operators to introduce a greater degree of programmability for purposes of service differentiation (such as QoS) and will eventually be used to handle large usage scenarios such as massive IoT. As end-end network orchestration solutions mature, network slicing will be able to handle inter-relationships between disparate network topologies, and multiple business interfaces in an automated and transparent manner.

Network slicing and 3GPP Activities

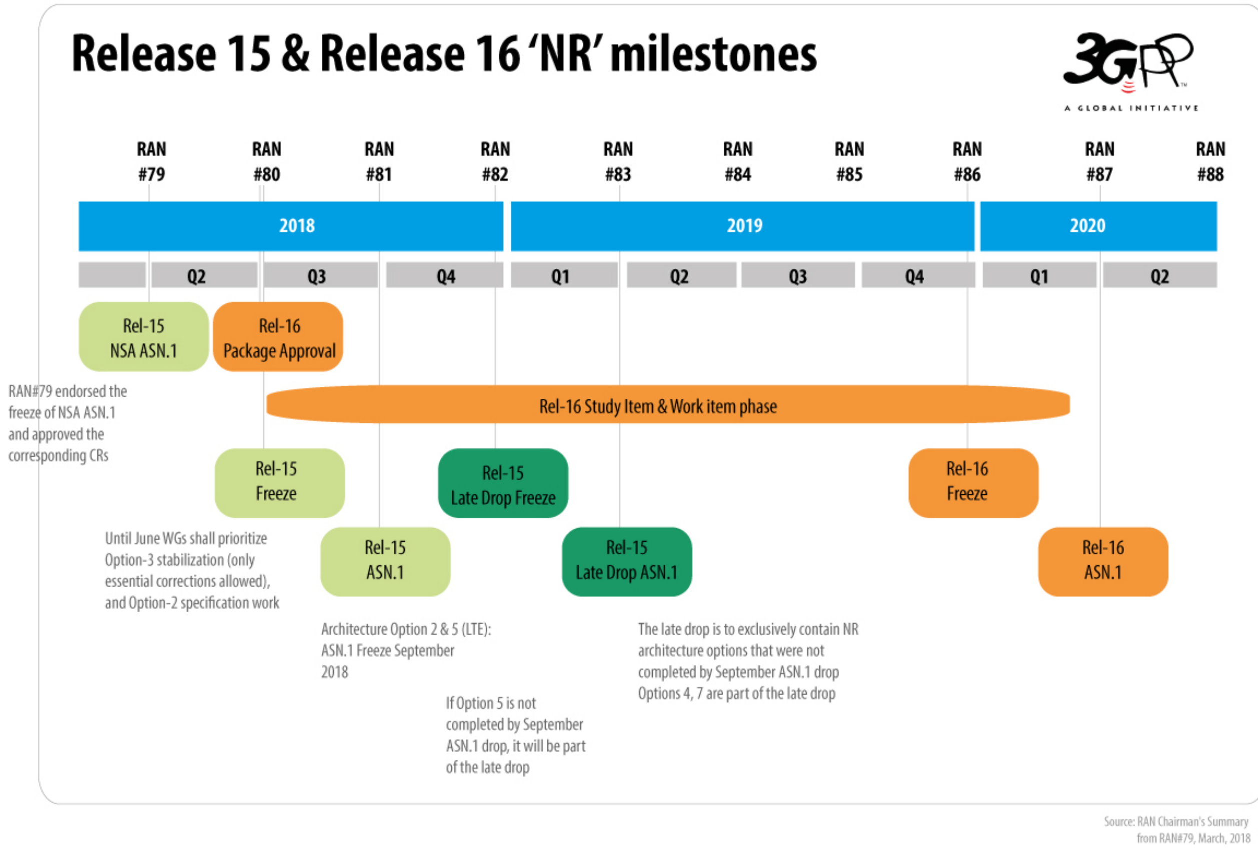


Figure 5: Source: RAN Chairman's Summary from RAN#79, March 2018

It is worth noting that 3GPP recognizes that the applicability of slicing extends beyond wireless contexts. In addition to the above Release 16 studies on slicing, 3GPP is also studying the overall "Wireless and Wireline Convergence for the 5G system architecture" in TR 23.716. The purpose of this study is to understand how the 5G Core network can be extended for wireline access usage (i.e., fixed access network defined by the Broadband Forum (BBF) and DOCSIS by CableLabs). How the 3GPP 5G QoS and network slicing frameworks can be utilized-by or extended-to the fixed access networks will also be part of this study.

3GPP Network Slicing requirements are included in 3GPP TS 22.261, Service requirements for the 5G system Stage 1, for Release 15 and updated for Release 16.

As defined by 3GPP, Network slicing allows the operator to provide customized networks. For example, there can be different requirements on functionality (e.g., priority, charging, policy control, security, and mobility), differences in performance requirements (e.g., latency, mobility, availability, reliability and data rates), or they can serve only specific users (e.g., MPS users, Public Safety users, corporate customers, roamers, or hosting an MVNO).

The 3GPP requirements are related to:

- Provisioning: create/modify/delete Network Slices, provision network functions to be used in Network Slices, define network services and capabilities supported by a network slice.
- Managing association to slices: configure association of devices and services to Network Slices, move/remove user between/from slices.
- Interoperating: support roaming and non-roaming using the same home slice, support devices simultaneously connected to multiple slices.
- Supporting performance and isolation: support dynamic slice elasticity, ensure performance isolation during normal and elastic slice operation and during slice creation or deletion, enable operators to differentiate performance and functionalities *between slices.*

A Network Slice is a complete logical network including Radio Access Network (RAN) and Core Network (CN). It provides telecommunication services and network capabilities, which may vary (or not) from slice to slice. Distinct RAN and Core Network Slices will exist. A device may access multiple Network Slices simultaneously through a single RAN.

3GPP TR 28.801, Study on management and orchestration of network slicing for next generation network, was completed for Release 15. This document specifies use cases for management of network slicing, potential requirements, potential solutions and recommendations.

3GPP defines 3 layers for network slicing:

1. Service Instance Layer,
2. Network Slice Instance Layer, and
3. Resource layer where each layer requires management functions, defined in 3GPP TR 23.799

The Study on management and orchestration of network slicing for next generation network investigates and makes recommendations on management and orchestration for network slicing on the Network Slice Instance Layer and for non-virtualized NE's also the Resource layer:

1. Use cases and requirements for management and orchestration of network slicing
2. Management and orchestration terminology and concepts for slices. The relationship between network slice management and orchestration concepts developed in the present document and the management and orchestration concepts defined by ETSI NFV.
3. Management of network function sharing in the context of network slicing (e.g. network function sharing support, and lifecycle management functions of the mobile network resources for network slices).
4. Impacts to management when a slice instance is shared between multiple parties (e.g. multiple partners and multiple domains etc.) based on the slice sharing, from 3GPP TR 23.799.
5. Isolation of management data between different parties within a slice instance if needed.
6. Impacts derived from Management of 3GPP Network Slices on the ETSI MANO architecture and procedures.
7. Automation of management and orchestration of network slice instances and the related policy configurations.
8. Management and orchestration mechanisms to support the isolation/separation of mobile network resources used by different network slice instances and the corresponding configuration of isolation/separation.
9. Solution for management and orchestration of network slicing and how it affects the specifications (e.g. Interface IRPs, Network Resource models and trace specifications).

Technical deep dive on slicing

5G Americas report on Network slicing provides examples of network slices including:

- A slice serving a utility company
- A slice serving remote control for a factory
- A slice serving a virtual operator
- A slice optimized for streaming video

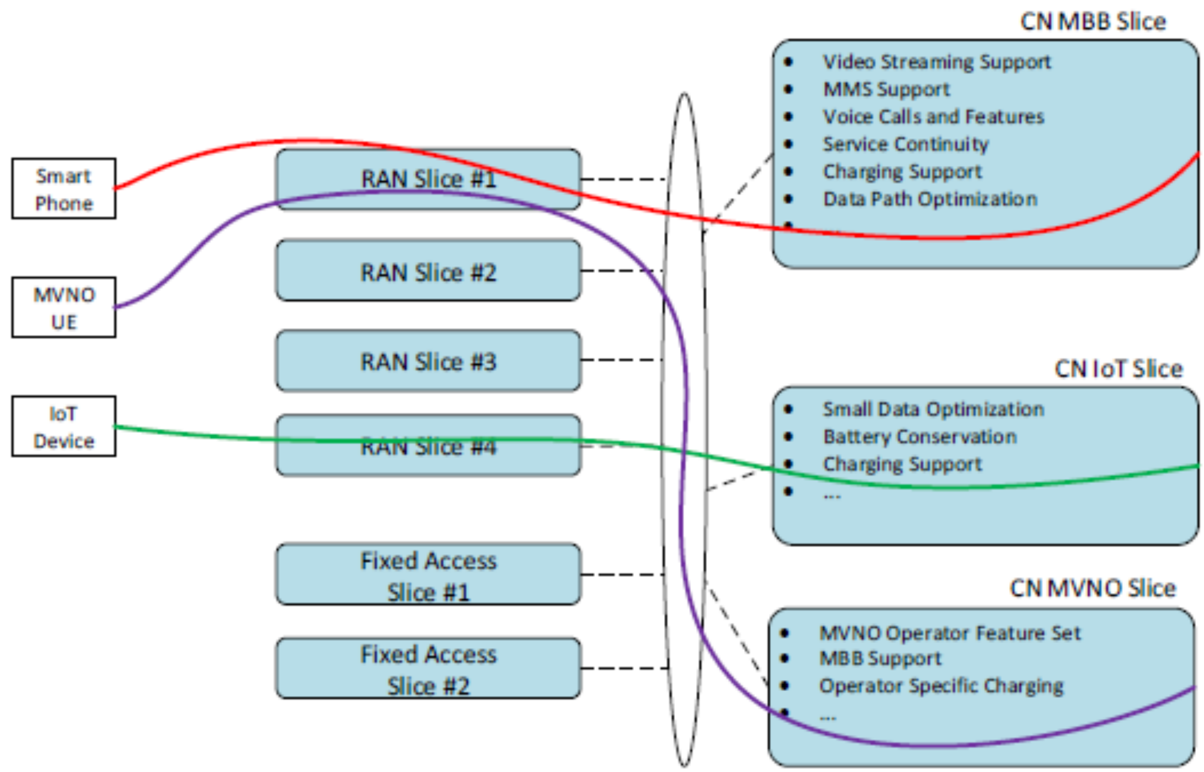


Figure 6

A Network Slice Instance (NSI) is a managed entity in the operator's network with a lifecycle independent of the lifecycle of the service instance(s). In particular, service instances are not necessarily active through the whole duration of the run-time phase of the supporting NSI. The NSI lifecycle typically includes an instantiation, configuration and activation phase, a run-time phase and a decommissioning phase. During the NSI lifecycle the operator manages the NSI.

The following phases describe the network slice lifecycle:

- Preparation phase
- Instantiation, Configuration and Activation phase
- Run-time phase
- Decommissioning phase

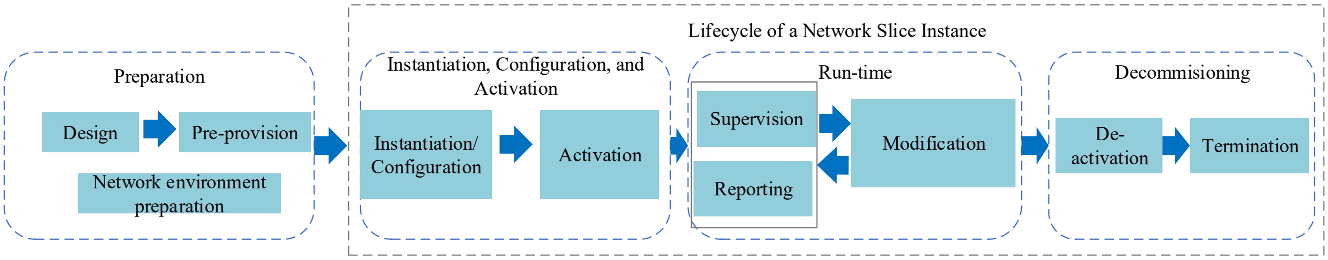


Figure 7: Lifecycle phases of an NSI

In the context of next generation networks, responsibilities regarding operations have to be clearly defined and assigned to roles. High-level business roles include:

- Communication Service Customer (CSC): Uses communication services.
- Communication Service Provider (CSP): Provides communication services Designs, builds and operates its communication services.
- Network Operator (NOP): Provides network services. Designs, builds and operates its networks to offer such services.
- Virtualization Infrastructure Service Provider (VISP): Provides virtualized infrastructure services. Designs, builds and operates its virtualization infrastructure(s). Virtualization Infrastructure Service Providers may also offer their virtualized infrastructure services to other types of customers including to Communication Service Providers directly, i.e. without going through the Network Operator.
- Data Center Service Provider (DCSP): Provides data center services. Designs, builds and operates its data centers.
- Network Equipment Provider (NEP): Supplies network equipment. For sake of simplicity, VNF Supplier is considered here as a type of Network Equipment Provider.
- NFVI Supplier: Supplies network function virtualization infrastructure to its customers.

- Hardware Supplier: *Supplies hardware.*

Depending on actual deployments, the roles defined above can be played by one single organization or by several ones. Hence, an organization can play one or several roles defined above.

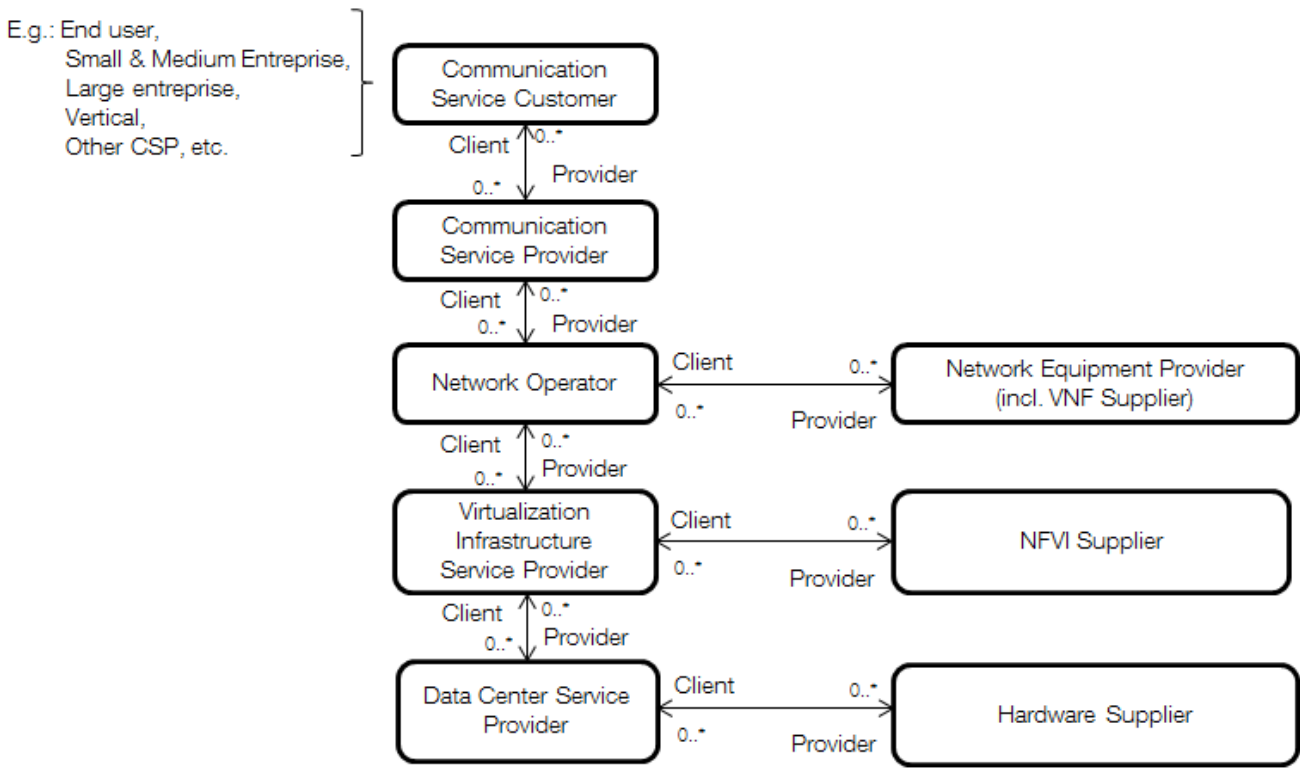


Figure 8:

Communication services offered by Communication Service Providers to Communication Service Customers are of various categories, among which:

- B2C services, e.g. mobile web browsing, VoLTE, Rich Communication Services, etc.
- B2B services, e.g. Internet access, LAN interconnection, etc.
- B2B2X services: e.g. services offered to other Communication Service Providers (e.g. international roaming, RAN sharing, etc.) or to Verticals (e.g. eMBB, etc.) offering themselves communication services to their own customers. B2B2X service type includes B2B2 applied recursively, i.e. B2B2B, B2B2B2B, etc.

A communication service offered by Communication Service Providers can include a bundle of specific B2C, B2B or B2B2X type of services. Taking as an example the B2C type of services, a bundle could include: data (for mobile web browsing), voice (through VoLTE), and messaging (via Rich Communication Services). In this case, each one of the individual B2C may be fulfilled by different PDU connectivity services provided via corresponding PDU sessions. According to 3GPP TR 23.501, a specific PDU session makes use of a single network slice, and different PDU sessions may belong to different network slices.

Depending on CSP requirements and/or Network Operator own policies, the Network Operator can decide to provide required network services using network slices. In such a case, only the Network Operator has knowledge of the existence of network slices; the CSP and its customers have not.

Network slice as a Service can be offered by CSPs to their customers, leaving the latter ones the possibility to offer their own services on top on network slice services (B2B2X). Such network slice services could be characterized by e.g.

- radio access technology,
- bandwidth,
- end-to-end latency,
- guaranteed / non-guaranteed QoS,
- security level, etc.

As an example, Multimedia Priority Service (MPS) can benefit from network slice as a Service. MPS as defined in 3GPP for LTE/4G (TS 23.401) allows certain subscribers (Service Users) priority access to system resources in situations such as congestion, creating the ability to deliver or complete sessions of a high priority nature. These Service Users can be government-authorized personnel, emergency management officials and/or other authorized users. MPS supports priority sessions on an "end-to-end" priority basis.

Since network slicing is not use for MPS today, Service Users can be rejected due to high traffic volume generated by non-Service Users. If ongoing traffic within the network cannot be preempted by the network processing node(s), Server Users must wait for the next system resources to be available. This is normally done by prioritizing the Service User at the system resource waiting queue in order to improve the establishment success rate but is not sufficient for first responders.

In 5G, a network slice can be created for Service Users of MPS. This slice uses dedicated network resources for MPS and is isolated from the network resources utilized by non-Service Users. During a high network load situation, Service users will not get their network resource blocked due to the activities caused by non-Service users. Using network slice as a Service will lessen the network congestion occurrences and improve the overall service experiences for Service Users of MPS.

Depending on service offering, the CSP imposes limits on the network slice characteristics that will be exposed to the CSC, and the CSP imposes limits on the network slice management capabilities that will be exposed to the CSC.

A CSP offering Network Slices as a Service relies on network services offered by one or more network Operators.

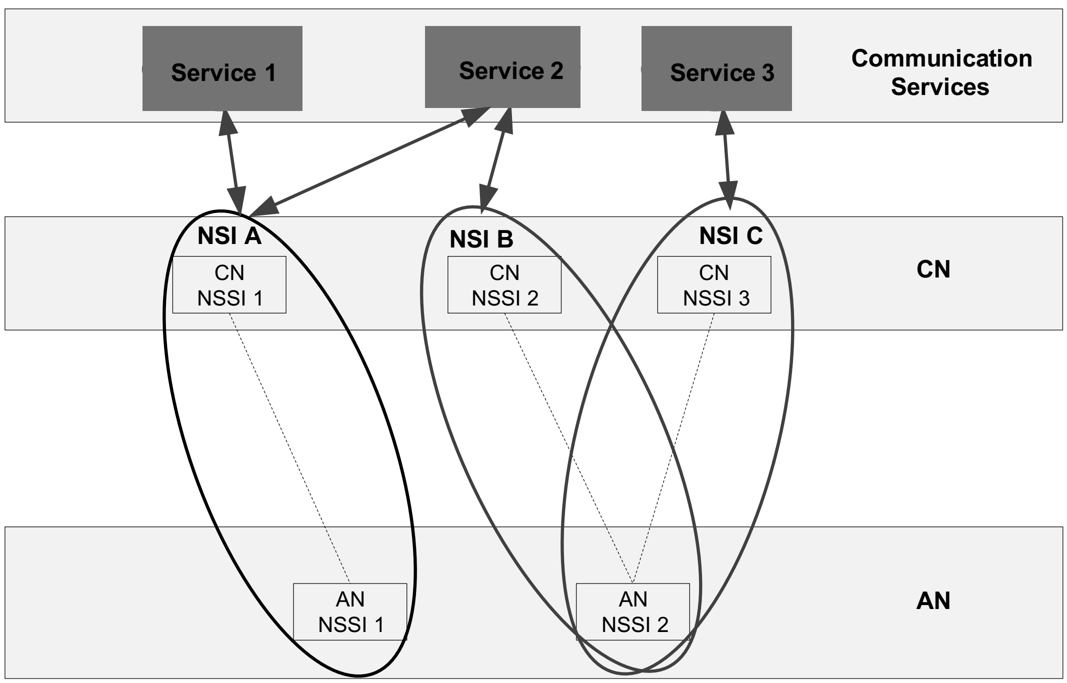


Figure 9

The following management functions are needed to manage the NSIs to support communication services:

- Communication Service Management Function (CSMF):
- Responsible for translating the communication service related requirement to network slice related requirements.

- Communicate with Network Slice Management Function (NSMF).
- Network Slice Management Function (NSMF):
- Responsible for management and orchestration of NSI.
- Derive network slice subnet related requirements from network slice related requirements.
- Communicate with the Network Slice Subnet Management Function (NSSMF) and Communication Service Management Function.
- Network Slice Subnet Management Function (NSSMF):
- Responsible for management and orchestration of NSSI.
- Communicate with the NSMF.

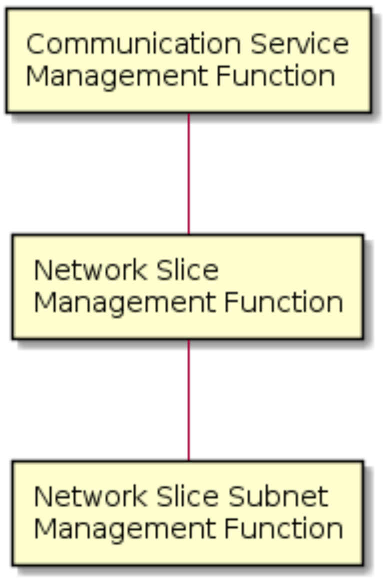


Figure 10

3GPP defines the following use cases for network slicing from the perspective of management of slices:

Use cases related to Network Slice management:

1. Network Slice Instance(s) lifecycle management
    a. Create a Network Slice Instance
    b. Create a Network Slice Instance with shared and non-shared CN network functions
    c. Create Network Slice Instance with shared Network Slice Subnet Instances
    d. Activate a Network Slice Instance

   e. De-activate a Network Slice Instance
   f. Modify a Network Slice Instance with shared and non-shared CN network functions and a shared AN
   g. Terminate a Network Slice Instance with shared and non-shared CN network functions and shared AN
   h. Change capacity of a Network Slice Instance
   i. Instantiate, configure and activate a Network Slice Instance in a network with virtualized NFs
2. Network Slice Instance(s) fault management
3. Network Slice Instance(s) performance management
4. Network Slice Instance(s) configuration management
5. Network Slice Instance(s) policy management
6. Communication service support with network slice
   a. Management support for network slicing
   b. Management support to provide a customer's service request using a Network Slice Instance
   c. Provide network slice as a service with guaranteed quality of service
7. Management data isolation
8. Multiple operator coordination management (roaming)
9. Automation of the network slice management

5G Americas describes the mapping among devices, access slices and CN slices can be 1:1:1 or 1:M:N, as the following illustrates (for example, a device could use multiple access slices, and an access slice could connect to multiple CN slices)

Services related to slicing

By definition, network slicing is intended to provide service differentiation. Mobile services in the 5G era can be highly specific to the level of specific applications and even industrial processes. These may require critical levels of reliability and low latency for example. Regulations defined for best effort mobile broadband may not have anticipated these sorts of applications. Operators should be able to define network slices to meet market needs and offer new services without fear that they will run afoul of regulations.

The GSMA has done some thinking on regulatory aspects of slicing, which can be found in reference 4 (network slicing use case requirements.) There is also some worry about European data protection / GDPR cross-border data transfers and slicing in the context of various industries.

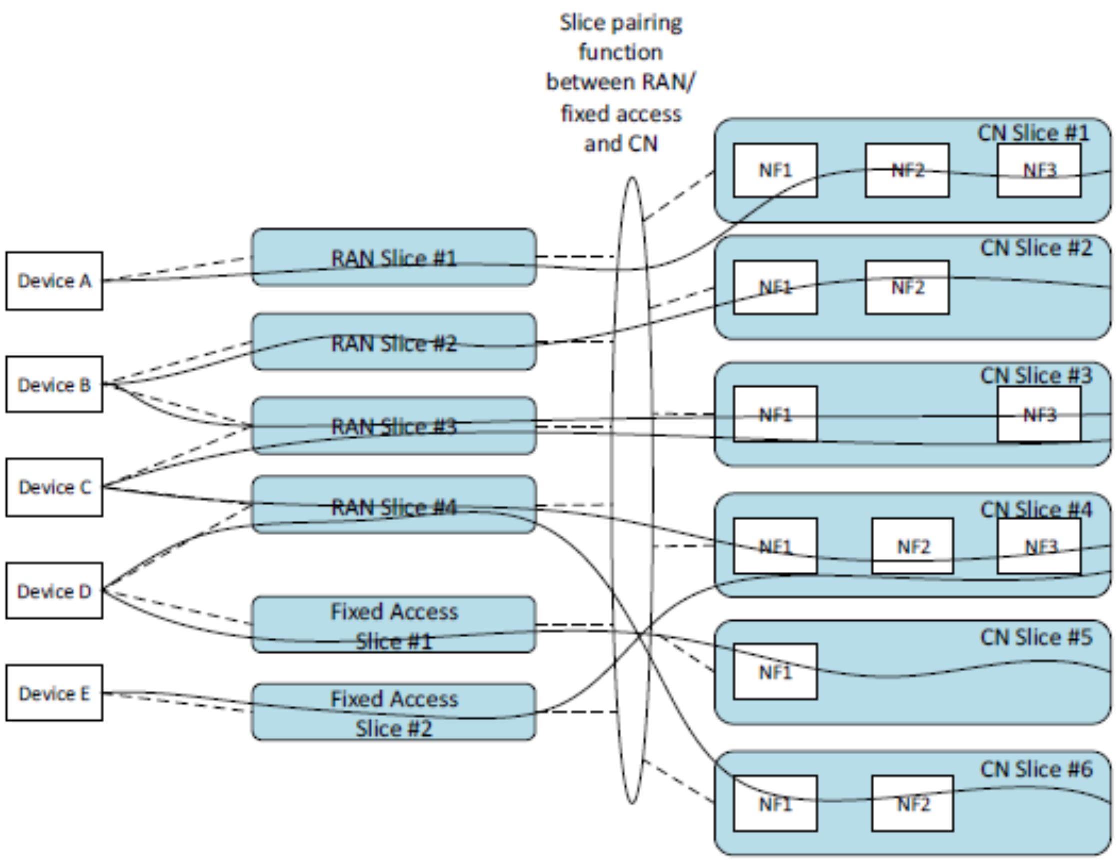


Figure 11

Summary

The multi-faceted capabilities that slicing brings to 5G will provide benefits to a wide range of recipients. Slicing will be used for network management by the operators, for new differentiated services, and for B2C business relationships. Some of the key points:

- Slicing provides the capability for differentiated, new service offerings
- Network slicing opens up new business models. This includes operator management beneficial capabilities, support for specific traffic parameters based on a particular vertical, operator to operator, MVN use of slices, emergency services, and more.
- Adoption may start small but is expected to grow as 5G is deployed and use cases develop.
- Involves different performance related SLAs: it is important that regulations do not inhibit adoption
- Slicing will enable the expansion of capabilities that can be integrated with enterprise VPN services, such as ultra-low latency localized hosting of enterprise/industrial applications
- Slicing is a refinement/evolution of virtual network concepts that have been used in wireline and wireless (per intro)
- 3GPP will continue to focus on slicing technology to support additional features and connections with non-5G slices
- Multi-operator slicing is unlikely to happen in the near term
- **Recommendation**: The FCC 5G/IoT working group has made a formal recommendation for the FCC to take action on providing “regulatory clarity” as it relates to slicing services. (assuming approval, paper will be published at fcc.gov/tac-reports-and-papers) Lack of clarity may hamper investments and or limit innovation related to new services.

Slicing Whitepapers to Reference:

1. *5G Americas: Network Slicing for 5G Networks & Services (34 pgs)* *http://www.5gamericas.org/files/1414/8052/9095/5G_Americas_Network_Slicing_11.21_Final.pdf*

2. *Small Cell Forum: URLLC Network Slicing (43pgs)* *https://scf.io/en/documents/199_-_URLLC_and_slicing_in_5G_small_cell_networks.php*

3. *GSMA - Introduction to Network Slicing (18pgs)* *https://www.gsma.com/futurenetworks/5g/introduction-to-5g-network-slicing/*

4. *GSMA - Network Slicing Use Case Requirements (54pgs)* *https://www.gsma.com/futurenetworks/5g/network-slicing-use-case-requirements-whitepaper/*

Glossary - Definitions

Term	Description
2G	Second Generation Mobile Network
3G	Third Generation Mobile Network GSM
3GPP	Third Generation Partnership Project
4G	Fourth Generation Mobile Network
5G	Fifth Generation Mobile Network
5GC	5G Core Network
5gNB	Fifth Generation NodeB
5GPPP	Fifth Generation Private Public Partnership
5GS	Fifth Generation System IMT
AMF	Access & Mobility Management Function
AP	Access Point
API	Application Program Interface
ARPANET	Advanced Research Projects Agency Network
AS	Application Server
ASN.1	Abstract Syntax Notation One
B2B	Business to Business
B2C	Business to Consumer
BBF	Broadband Forum
BIAS	Broadband Internet Access Service
CA	Carrier Aggregation
CAPEX	Capital Expenditure
CN	Core Network
CP	Control Plane
CPE	Customer Premise Equipment
CS	Circuit-Switched
CSC	Customer Service Customer
CSCF	Call Session Control Function
CSFB	Circuit Switched Fallback

CSMF	Communication Service Management Function
CSP	Communication Service Provider
CUPS	Control Plane – User Plane Separation
DARPA	Defense Advanced Research Projects Agency
DCSP	Data Center Service Provider
DOCSIS	Data Over Cable Service Interface Specification
DSL	Digital Subscriber Line
E-UTRA	Evolved Universal Terrestrial Radio Access
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
E2E	End-to-End
eDÉCORE	Enhancements of Dedicated Core Networks
eMBB	Enhanced Mobile Broadband
EMM	EPS Mobility Management
EN-DC	E-UTRA-NR Dual Connectivity
eNB	Evolved Node B
EPC	Evolved Packet Core
ePDG	Evolved Packet Data Gateway
EPS	Evolved Packet System
ESM	EPS Session Management
ETSI	European Telecommunications Standards Institute
EU	European Union
EVS	Enhanced Voice Services
FCC	Federal Communications Commission
FTTH	Fiber to the Home
GDPR	General Data Protection Regulation
Global	System for Mobile Communications
GSMA	Groupe Spéciale Mobile Association
GW	Gateway
HEVC	High Efficiency Video Coding
HFC	Hybrid-Fiber Coax

HO	Handover
HSS	Home Subscriber Server
IMS	IP Multimedia System
IMT	International Mobile Telecommunications
IoT	Internet of Things
IP	Internet Protocol
IPTV	Internet Protocol Television
ISD	Inter-Site Distance
ISP	Internet Service Provider
ITU	International Telecommunication Union
LCS	Location Services
LAN	Local Area Network
LIPA	Local IP Access
LS	Liaison Statement
LTE	Long Term Evolution
MCG	Master Cell Group
MCPTT	Mission Critical Push-To-Talk
mIoT	Massive IoT
MME	Mobility Management Entity
mMTC	Mobile Machine Type Communications MN Master Node
MPS	Multimedia Priority Service
MR-DC	Multi-RAT Dual Connectivity
MVNO	Mobile Virtual Network Operator
NAS	Non-Access Stratum
NB-IoT	Narrowband-IoT
NCPs	Network Control Points
NE	Network Element
NE-DC	NR E-UTRA Dual Connectivity
NEP	Network Equipment Provider
NEST	Network Slicing Templates

NF	Network Function
NFV	Network Function Virtualization
NFVI	Network Function Virtualization Infrastructure
NOP	Network Operator
NRF	Network Repository Function
NSA	Non-Standalone
NSF	National Science Foundation
NSFnet	National Science Foundation Network
NSI	Network Slice Instance
NSMF	Network Slice Management Function
NSSI	Network Slice Subnet Instance
NSSMF	Network Slice Subnet Management Function
NWDA	Network Data Analytics
OAMP	Operations, Administration, Maintenance, Provisioning
OEM	Original Equipment Manufacturer
OTT	Over the Top
P-CSCF	Proxy Call Session Control Function
PCF	Policy Control Function
PCRF	Policy and Charging Rules Function
PDCP	Packet Data Convergence Protocol
PDU	Protocol Data Unit
PGW	Gateway PGW-C Control plane of the PGW
PGW-U	User plane of the PGW
PON	Passive-Optical Network
PSTN	Public Switched Telephone Network
QCI	QoS Class Identifier
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RCS	Rich Communication Services

RRC	Radio Resource Control
RX	Reception
SA	Standalone
SBA	Service Based Architecture
SCG	Secondary Cell Group
SDN	Software-Defined Network
SGW	Serving Gateway
SIG	Special Interest Group
SIP	Session Initiation Protocol
SIPTO	Selected IP Traffic Offload
SLA	Service Level Agreement
SM	Session Management
SMF	Session Management Function
SMS	Short Message Service
SN	Secondary Node
SP	Service Provider
SRVCC	Single Radio Voice Call Continuity
SST	Slice/Service Types
TB	Terabyte
TCO	Total Cost of Ownership
TDM	Time Division Multiplexing
TE	Terminal Equipment
TR	Technical Requirement
TS	Technical Standard
TV	Television
TWAG	Trusted Wireless Access Gateway
TX	Transmission
UDM	User Data Management
UE	User Equipment
UL	Uplink
Ultra-Reliable	Low Latency Communications

UMTS	Universal Mobile Telecommunications System
UPF	User Plane Function
URLLC	Ultra-Reliable Low Latency Communications
USA	United States of America
V2X	Vehicle to Vehicle or Infrastructure
ViLTE	Video over LTE
VISP	Virtualization Infrastructure Service Provider
VLAN	Virtual Local Area Network
VNF	Virtual Network Function
VoLTE	Voice over Long-Term Evolution
VoWiFi	Voice over WiFi
VPLS	Virtual Prlvate LAN Service
VPN	Virtual Private Network
WLAN	Wireless Local Area Network