



5G Service-Guaranteed Network Slicing White Paper



Issue V1.0
Date 2017-02-28

Abstract

Previous generations of mobile networks enabled voice, data, video, and other life-changing services. In comparison, 5G will change our society by opening up the telecom ecosystem to vertical industries. 5G will help vertical industries to achieve the “Internet of Everything” vision of ubiquitously connected, highly reliable, ultra-low latency services for massive number of devices. Service-guaranteed network slicing introduced in this white paper is one of the essential features for 5G to achieve this vision. Key players from operators, vendors, and vertical industries have come together to establish a common understanding on service-guaranteed network slicing in terms of the vision, end-to end (E2E) solution, key enabling technologies, and the impacts for vertical industries. This white paper describes the thinking on network slicing in 5G.

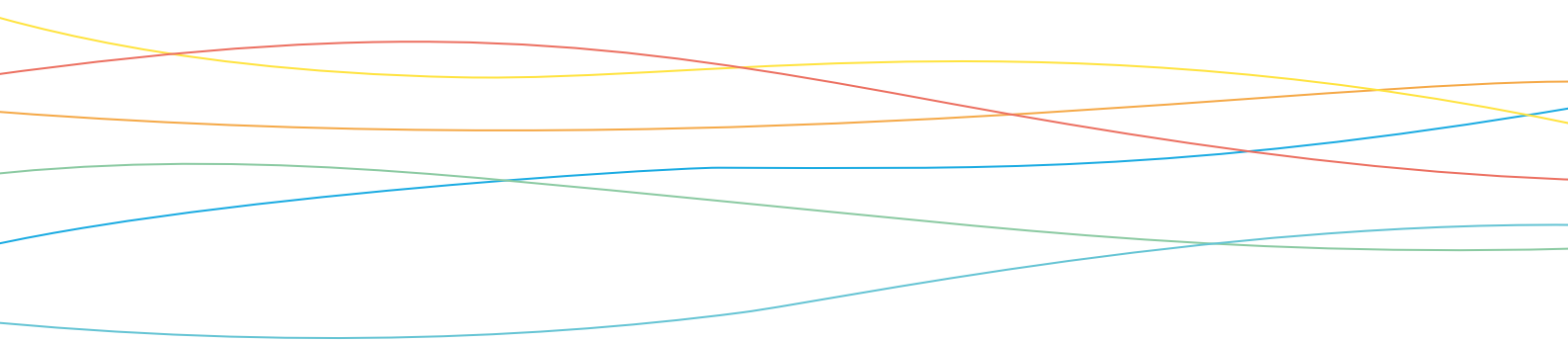
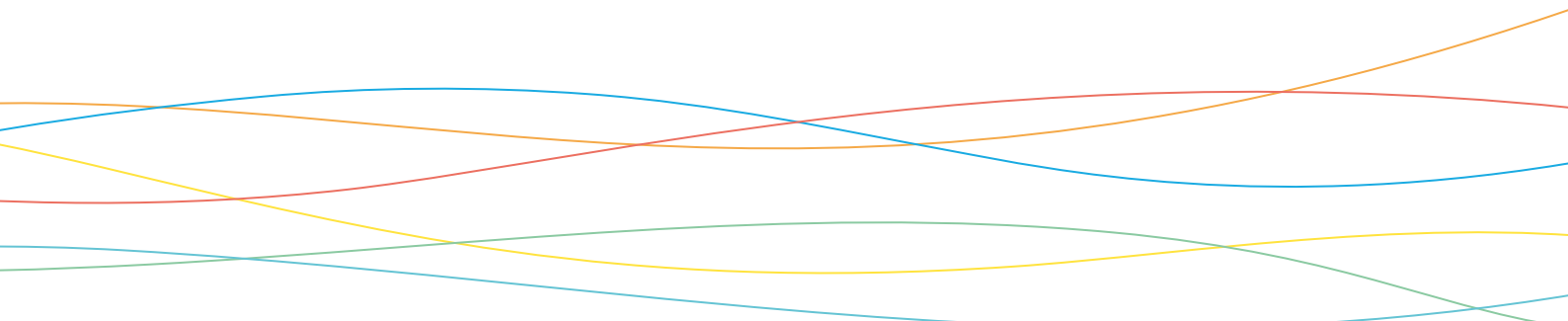




Table of Contents

1. Industry Trends and Requirements	02
2. Visions of Service-Guaranteed Network Slicing	04
3. Overall Architecture of Service-guaranteed Network Slicing	06
3.1 Concepts	06
3.2 Concept Clarifications	06
3.3 Architecture	07
4. Key Technologies to Enable Service-Guaranteed Network Slicing	09
4.1 Network Management System	09
4.1.1 Network Slice Management (NSM) Architecture	09
4.1.2 Network Capability Exposure via Business Support System	10
4.1.3 Third-party Applications	11
4.2 Security	11
4.2.1 Infrastructure Security	11
4.2.2 Network Management Security	11
4.2.3 NSI Security	12
4.3 Enabling Technologies for Different Technical Domains	12
4.3.1 Access Network	12
4.3.2 Core Network	14
4.3.3 Transport Network	15
4.3.4 Terminal	17
4.4 Technology Evolution	17
5. Use Case for Service-Guaranteed Network Slicing	18
6. Summary and Suggestions	20



1. Industry Trends and Requirements

The 5G networks are not only envisioned as a support for “Internet of Things” (IoT), but also as means to give rise to an unprecedented scale of emerging industries, instilling an infinite vitality in future telecommunications. IoT requires support for a diverse range of service types, such as eHealth, Internet of Vehicles (IoV), smart households, industrial control, environment monitoring, and so on. These services will drive the rapid growth of IoT and facilitate hundreds of billions of devices to connect to the network, which also conceives the “Internet of Everything” vision especially from vertical industries.

The requirements for IoT services are also very

diverse. Services such as smart households, smart grid, smart agriculture, and intelligent meter reading, will require supporting an extremely large number of connections and frequently transmitted small data packets. Services such as smart vehicles and industrial control will require millisecond-level latency and nearly 100% reliability, while infotainment services will require extreme fix/mobile broadband connectivity. These requirements indicate that the 5G networks need be more flexible and scalable to support massive connections of different nature. Meanwhile, operators will perform a gradual shift away from pipe services towards coping with vertical industry needs:

- **Service diversity**

The services foreseen in the 5G era fall into three typical scenarios: enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low Latency Communications (URLLC), and massive Machine Type Communications (mMTC). eMBB focuses on services characterized by high data rates, such as high definition (HD) videos, virtual reality (VR), augmented reality (AR), and fixed mobile convergence (FMC). URLLC focuses on latency-sensitive services, such as self-driving, remote surgery, or drone control. mMTC focuses on services that have high requirements for connection density, such as those typical for smart city and smart agriculture use cases. Each scenario requires a completely different network service and poses requirements that are radically different, sometimes even contradictory.

- **Guaranteed performance**

Several key performance indicators (KPIs) must be simultaneously satisfied for some of the above-mentioned services. For example, VR and AR have strict requirements on data rate as well as latency. Such demands become more stringent for vertical industries, where the terminals are normally "machines" with very low tolerance on performance degradation.

- **Fast deployment and short time-to-market (TTM)**

It is a long process to deploy conventional mobile networks. A simple service update may take from 10 to 18 months. Such long cycles are very difficult to meet tailored and fast service provisioning and short TTM demands from vertical industries.

- **Resource multiplexing and isolation**

Different from current telecom practice, vertical industries are likely to get involved with specialized network functions (dedicated routing, mobility support, customized flow handling, in-network processing, etc.). To handle such diversity without losing operation efficiency, operators prefer to use resource multiplexing approach with secured isolation provisioning.

- **Automation**

Flexibility and scalability are the key features of the 5G networks. Such networks cannot depend on manual management. Fully automatic network management techniques, such as self-diagnosis, self-healing, automatic configuration, self-optimization, and auto installation/plug-and-play, are fundamental to achieve efficient network operations and to provide the dynamic service mix. With the progress of the automatic network management techniques, management will become more agile and more adaptive. New tools for such management are required; in particular, artificial intelligence (AI) and automatic learning techniques should be considered for the 5G networks.

- **New ecosystem and business model**

The 5G networks will support new roles and business models, which may involve network infrastructure providers, operators (mobile network operators, mobile virtual network operators, etc.), and vertical service providers. These new roles and business relationships help the telecom industry to build a new ecosystem together with vertical industries.

- **Convergence of fixed and mobile access**

FMC is also a very important requirement, because customers do expect the same user experience regardless of the access technology used. While today the architectures, service concepts and ecosystems of fixed and mobile networks differ in many aspects, it is envisioned that with 5G these will converge. An architecture that can natively handle all kinds of fixed and mobile access technologies will contribute significantly to enable the design goal of truly converged 5G networks.

2. Visions of Service-Guaranteed Network Slicing

In the 5G era, vertical industries will trigger the networks to shift from the traditional “human-centric” services to “machine-centric” services. This not only allows the

telecommunication industry to develop a new ecosystem, but also becomes the new engine to boost the social economy with the following core visions:

Vision 1: Provide guaranteed performance to meet the fundamental service requirements of vertical industries.

Upon the fundamental connectivity service, guaranteed performance (e.g., latency, data rate, reliability, connectivity, and power consumption) will enable operators to embrace vertical industries in 5G ecosystem. Guaranteed performance is not only about quality of service (QoS), it also implies customized network functions and resources to tackle different types of services, for instance, to provide vehicle-to-everything (V2X) service with customized mobility management.

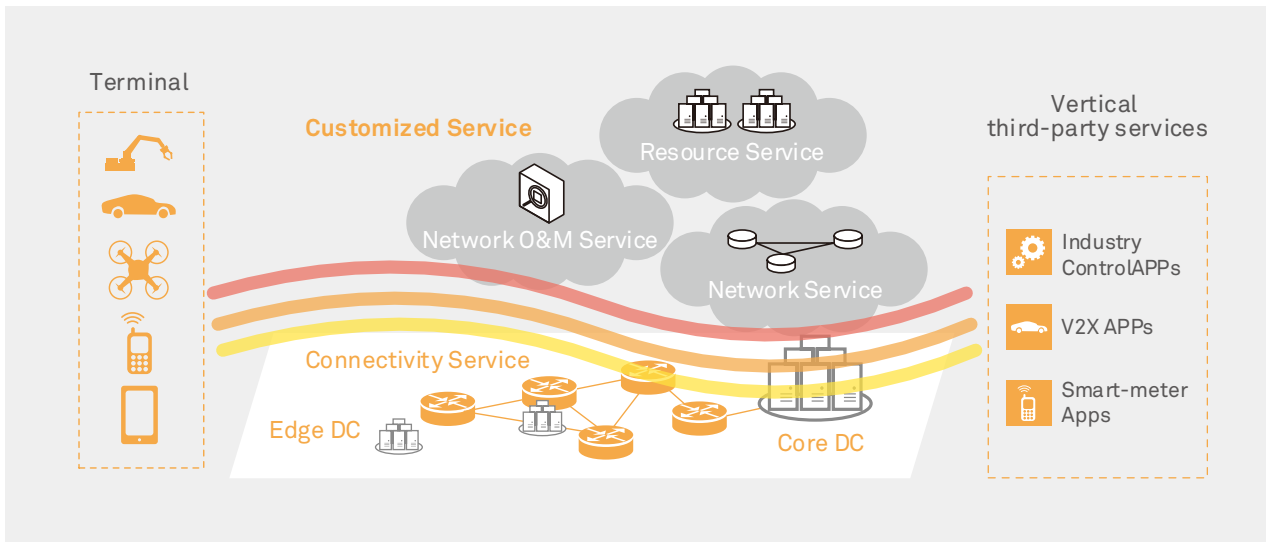
Vision 2: Provide customized services to enhance the competence of vertical industries.

Provisioning a guaranteed performance is only the basic proposition to cooperate with vertical industries in 5G. The further essential step towards success is to bring more concrete value for the vertical services, for instance, reducing their service operational cost and capital cost, shortening TTM, etc. Helping vertical industries to increase their competence is a vital component of

the 5G ecosystem.

Based on the fundamental connectivity services, operators should invoke deeper business potentials via providing customized services, for instance:

- Network services: The network capabilities, e.g., caching, can be used to enhance vertical service performance.
- Resource services: Vertical industries are encouraged to deploy their services in the operator’s edge data centers (DCs) and core DCs, because operators could use the advantage of the orchestration of network and cloud resource, as well as edge computing.
- Network operation and maintenance (O&M) services: Independent O&M according to customized policies is an appealing feature for vertical industries.



· Figure 1: Service-guaranteed network slicing vision

As presented above, the flexibility and diversity expectations from the core visions are real and tremendous. The question is how to fulfill these: the flexibility of services on the one hand and the diversity of necessary network technologies on the other hand pose a daunting requirement on the network design, control, operations and management. Such a system bears a high risk of crumbling under its own complexity. To overcome these challenges while still fulfilling the expected future demands, a service-guaranteed network slicing is introduced in this white paper, aiming to realize the above core visions. It proposes to have

several logical networks with different network services, provisions, mechanisms, or assurances on the same infrastructure. Vertical industries interested in the supported services therefore would only be required to concentrate on the management of the network slicing specific provisions, tightly coupled with the expected services. Such concentration on the business needs ensures interest and competence of vertical industries on the one hand and, on the other hand, offload them from complex considerations of designing, deploying, testing and running such networks.

3. Overall Architecture of Service-guaranteed Network Slicing

3.1 Concepts

Since “network slicing” appeared in the 5G vocabulary, a number of concepts have been derived from it, i.e. network slicing instance, network slice type, etc. This section aims to clarify the definition of these concepts and their corresponding relationships:

- Network slicing: Network slicing is the collection of a set of technologies to create specialized, dedicated logical networks as a service (NaaS) in support of network service differentiation and meeting the diversified requirements from vertical industries. Through flexible and customized design of functions, isolation mechanisms, and O&M tools, network slicing is capable to provide logical dedicated networks upon a common infrastructure.
- Network slice instance (NSI): An NSI is the realization of network slicing concept. It is an E2E logical network, which comprises of a group of network functions, resources and connection relationships. An NSI typically covers multiple technical domains, which includes terminal, access network (AN), transport network (TN) and core network (CN), as well as DC domain that hosts third-party applications from vertical industries. Different NSIs may have different network functions and resources. They may also share some of the network functions and resources.
- Network slice type: Network slice types are high-level categories for NSIs, which reflect the distinct demands for network solutions. Three fundamental network slice types have been identified for 5G: eMBB, mMTC, and URLLC. These could be further extended, e.g. according to the operator’s policies or with the development of 5G.
- Network slice template: Network slice template is the output of the slice design phase used to create NSIs.
- Tenant: Tenants are the operators' customers (for example, customers from vertical industries) or the operators themselves. They utilize the NSIs to provide services to their users. Tenants typically will have independent O&M requirements, which are uniquely applicable to the NSIs.

3.2 Concept Clarifications

The aforementioned key concepts have the following relationships.

- Network slice types and tenants are important references for creating an NSI. An NSI is instantiated from one network slice template with a specific network slice type. A tenant that provides different service types may use multiple NSIs with different network slice types. For tenants, who may provide services of the same service type, they can still use differentiated NSIs via the

customization of the network slice template with the same network slice types.

- Network slice template design is separate from the NSI operation. In the design phase, the network slice template is generated based on the network capability of each technical domain and a tenant's particular requirements. In the operation phase, an NSI is instantiated based on the network slice template, which includes the deployment and configuration of related network functions and related resources in different technical domains. The network slice design is separate from the operation to enable the repeated use of a network slice template.
- NSIs require multi-dimensional management. An NSI usually includes multiple technical domains. An NSI may also include multiple administrative domains that belong to different operators. To guarantee NSI's fast deployment, it is essential to use efficient multi-dimensional management via coordination and cooperation across such different domains.
- NSIs ensure SLA compliance. Tenants will sign service-level agreement (SLA) with operators, which may include requirement agreements related to security/confidentiality, visibility/manageability, specific service characteristics (service type, air interface standard, and customized functions), and corresponding performance indicators (latency, throughput, packet loss rate, call drop rate, and reliability/availability).
- Terminals may be involved in the selection of NSIs. Terminals can access one or multiple NSIs. Terminals could assist NSI selection based on, for instance, network slice type, while the network performs the final selection decision. Simple terminals, such as sensors, are usually in a static and one-to-one relationship with NSIs, because the costs and power consumption requirements limit the terminal capability. Therefore, the NSI selection is solely performed by the network.

3.3 Architecture

Enabling network slicing in 5G requires native support from the overall system architecture. As shown in Figure 2, the overall architecture consists of three fundamental layers: the infrastructure layer, network slice layer and network management layer. The infrastructure layer provides the physical and virtualized resources, for instance, computing resource, storage resource, and connectivity. The network slice layer runs above the infrastructure layer and provides necessary network functions, tools and mechanisms to form end-to-end (E2E) logical networks via NSIs. The network management layer contains the generic BSS/OSS and network slice management (NSM) system, which designs and manages network slicing. Moreover, it also assures the SLA requirements.

The overall architecture has the following key features:

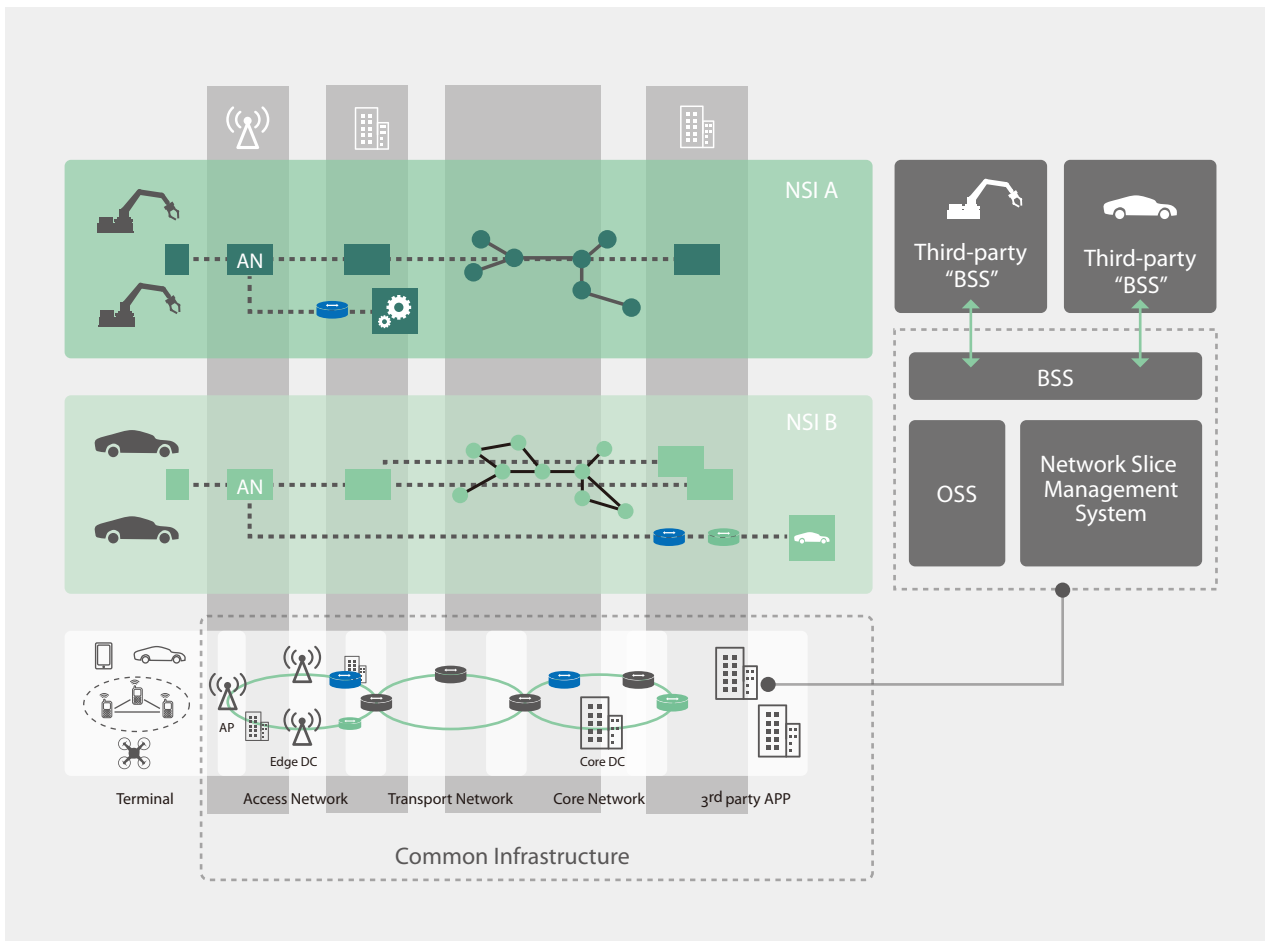
- Common infrastructure: Being different from the dedicated network solution that uses physically isolated and static networks to support tenants, network slicing promotes the use of a common infrastructure among tenants from the same operator. It helps to achieve higher resource utilization efficiency and reduce the service TTM. Moreover, such design is beneficial for long-term technology evolution as well as for shaping a healthy industry ecosystem.
- On-demand customization: Each technical domain in an NSI has different customization capabilities, which are coordinated through the NSM system during the process of network slice template design, and NSI deployment and O&M. Each technical domain can perform an independent tailoring-process in terms of design schemes to achieve an effective balance between the simplicity needed by commercial practice and architectural complexity.
- Isolation: The overall architecture supports the isolation of NSIs, including resource isolation, O&M isolation, and security isolation. NSIs

can be either physically or logically isolated at different levels.

- **Guaranteed-performance:** Network slicing seamlessly integrates different domains to meet and ensure industry-defined 5G performance specifications and to accommodate vertical industry requirements.
- **Scalability:** Due to virtualization, which is one of the key enabling technologies for network slicing, resources occupied by an NSI can dynamically change, e.g., scaling in/out.
- **O&M Capability Exposure:** Tenants may use dedicated, shared or partially shared NSIs.

Furthermore, different tenants may have independent O&M demands. The NSM system provides access to a number of O&M functions of NSIs for the tenants, which for instance allows them to configure NSIs related parameters, e.g., policy.

- **Support for multi-vendor and multi-operator scenarios:** Network slicing allows a single operator to manage multiple technical domains, which may be composed of network elements supplied by different vendors. In addition, the architecture also needs to support the scenario, where the services from the tenants may cover different administrative domains owned by different operators.



· Figure 2: Overall architecture to enable network slicing

4. Key Technologies to Enable Service-Guaranteed Network Slicing

4.1 Network Management System

4.1.1 Network Slice Management (NSM) Architecture

The NSM system plays an important role in the entire system architecture. It provides the following services:

- **Design:** design network slice templates according to the network capabilities and SLA requirements.
- **Provisioning:** comprise slice instantiation, configuration, and activation.
- **Runtime assurance:** observe the running status of NSIs and ensure SLA.
- **Decommissioning:** delete an NSI when its services are not used anymore.

The NSM shall be based on the state of the art cloud management technologies with enhanced features to support network slicing. It provides O&M capability using a streamline of aforementioned services, which address inadequacies of the traditional network management system, e.g., long TTM or lack of automatic O&M methods. The NSM system could further help operators to establish an open ecosystem to enable new business opportunities.

Figure 3 depicts the overall NSM system architecture, which uses “Layer- and Domain-based management” design principle. “Layer-based” management defines two layers within the NSM: slice support system (SSS) and domain slice support system (DSS). “Domain-based” management implies that the basic capabilities are provided by each individual technical domain. The cooperation between the DSS and SSS guarantees the E2E SLA.

• Slice Support System (SSS)

The SSS mainly comprises two functional blocks: the Network Slice Template Designer and the Cross Domain Slice Manager. The former generates the network slice template according to the network capability of each technical domain as well as the functional and performance requirements from the tenants. The latter is responsible for the NSI lifecycle management (i.e. provisioning, runtime assurance, and decommissioning). The SLA is guaranteed through multi-dimensional coordination among different domains. Based on the capability of each technical domain, the SSS decomposes an SLA in terms of sets of requirements and maps each segment of SLA to the corresponding technical domain. To ensure the overall SLA, the SSS aggregates the network service performance from each individual techni-

cal domain. Based on this, the SSS performs necessary adjustments and configurations to ensure closed-loop control.

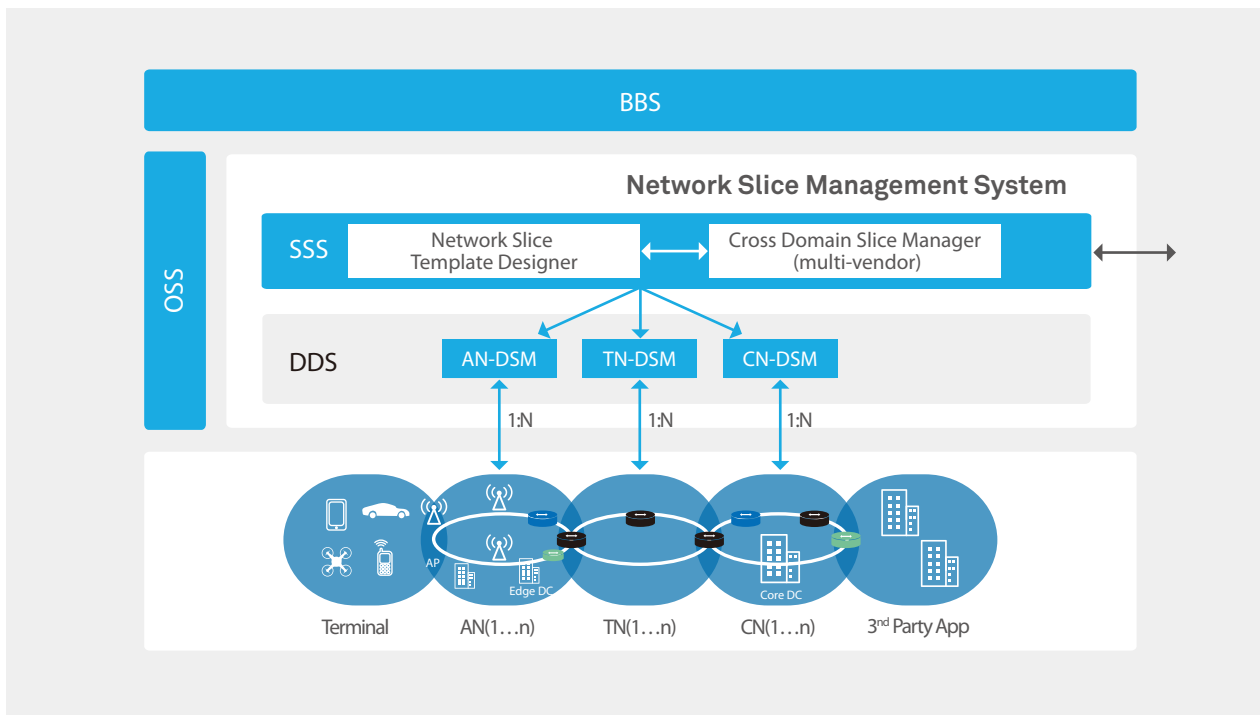
To support management across different administrative domains for different operators, the interworking between different SSSs is compulsory.

• **Domain slice Support System (DSS)**

The DSS comprises the Domain Slice Managers (DSMs) for different technical domains: access network DSM (AN-DSM), core network DSM (CN-DSM), and transport network DSM (TN-DSM). As a logical entity, the DSM is responsible for the design, provisioning, runtime assurance, and decommissioning of subnets in a single technical domain. The

DSS ensures the real-time guarantee for decomposed SLA capabilities in each domain, e.g. via monitoring and fault localization. Each domain has independent SLA-specific closed-loop control of functions and resources for fast service scheduling and resource optimization.

The task of the NSM system is not only about seamlessly managing and assuring the SLA, together with advanced AI algorithms, it could also predict the network status changes in order to provide certain management and control actions for precaution. The NSM system could be standalone (a new management entity) or non-standalone (integrated with OSS).



· Figure 3: Network slice management (NSM) architecture

4.1.2 Network Capability Exposure via Business Support System

The Business Support System (BSS) from operators is directly facing the tenants. Therefore, its usability is an essential factor. Operators use the BSS to provide their abstracted network capability to the tenants. It mainly supports the following

capabilities: design, purchasing, deployment, and monitoring.

• **Design** includes the design and offering of commercial products related to network slicing. Based on the service types and tenants' requirements, the SLA is formulated. A purchasable product may use one or more NSIs to accommodate

the tenants' service. Such product with packaged NSIs is used as an offering for the tenants, which focus on the commercial attributes of products, such as, pricing and sales territory.

- **Purchasing** is the key part for the user experience of the tenants. For instance, it is essential for the BSS to have a well-designed store page for displaying the products and personal center for monitoring the purchasing progress and triggering network service related upgrading process.
- **Deployment** of a product is triggered by the BSS after a successful customer purchase.
- **Monitoring** refers to the BSS capability of allowing tenants to view the operational as well as performance related information for the running services, e.g., throughput and latency of certain NSIs.

4.1.3 Third-party Applications

The flexibility and customization of network slicing are also reflected in the accommodation of third-party applications. In addition to the various network functions provided by operators, it is also feasible to deploy third-party applications on NSIs to meet the specific requirements from the tenants. Such third-party applications could be from tenants directly, or from non-tenant parties (e.g., tenants' customer or provider).

The main reason to support the deployment of third-party applications is to enable services with specific requirements, such as, URLLC services requiring ultra-low latency. It is beneficial to reduce the length of the transmission path by moving the network functions and third-party applications close to the AN, e.g., leveraging the advantage of edge computing.

In addition, third-party applications can also provide substitution of network functions, such as user-customized authentication and mobility management, which are designed especially to support their own services. Other than control plane related network functions, customized user plane network

functions, such as service gateway from tenants can be also deployed within the operator networks. This would enable preliminary filtration and aggregation of a large amount of data (e.g., from sensors). The NSM system should thus support the deployment of third-party applications. The deployment positions can be either specified, e.g., in an AN, CN, or dynamically determined by the SSS based on service requirements and network conditions during the network slice design phase.

4.2 Security

The overall architecture defined in the previous section contains three fundamental layers: the infrastructure, network slice, and network management layer. Each layer must consider its individual security risks and protection mechanisms. Moreover, it is necessary to consider all domains together as an organic whole to provide overall security. In general, there exist the following three aspects in a holistic framework of network slice security.

4.2.1 Infrastructure Security

As NSIs are sharing the same infrastructure, proper isolation between NSIs must be enforced to avoid adverse cross-effects and information leakage, especially when NFV is used. For example, different virtual machines or containers are used for different network functions and the virtual links connecting VNFs dedicated for different NSIs should be logically isolated.

4.2.2 Network Management Security

Security risks exist in every phase of the NSI lifecycle management in the network management layer. Malicious attacks may use malware to compromise a network slice template, threatening all subsequent NSIs. Attacks may also pass through configuration interfaces during the runtime phase of an NSI. Confidential data could be obtained during the decommissioning phase, if the NSI is handled improperly. Therefore, the security considerations should cover each single step of the lifecycle management of NSIs.

As some network capabilities and interfaces are exposed to tenants, the capabilities granted to a particular tenant are defined by the operator. Tenants must be authenticated and authorized before being allowed to access these capabilities and interfaces.

4.2.3 NSI Security

To guarantee security for the network services provided by an NSI, it requires embedding the security mechanism and security provisioning entity (e.g. security anchors and security functions) into the logical network architecture of the NSI.

Security isolation: Without security isolation, malicious attacks with access to one NSI may use that NSI as a launching pad for attacking other NSIs by, for instance, illegally occupying resources of another NSI. In addition, it may also result in breaches of data confidentiality and integrity attacks.

Slice access control: A terminal should be authenticated and authorized to access a specific NSI. The communication between the terminal and the allocated NSIs should be protected against attacks.

Customized security mechanisms: Terminals may require different levels of security protection. Terminals accessing the eMBB type NSI have strict security requirements on authentication and encryption/decryption, which can be similar to the mechanisms used in LTE. Terminals like low-cost sensors accessing the mMTC type NSI require only lightweight authentication and encryption/decryption algorithms due to limited computing capability. Terminals accessing the URLLC type NSI require quick access authentication and strong encryption algorithms.

4.3 Enabling Technologies for Different Technical Domains

Network slicing that empowers services with guaranteed-performance relies on the enabling technologies from all technical domains and E2E operation and management. The following sections illustrate the key technologies and architecture from the perspectives of the AN, CN, TN, and terminal.

4.3.1 Access Network

With the development of 5G, new radio access technology (RAT) is also emerging, which aims to bring higher access performance compared to the 4G RAT. The new AN architecture, which includes fixed and mobile access, is expected to provide native support for network slicing. Hence, it is essential to understand how to enable isolation, capability of customization, and guaranteed SLA in the AN, while concurrently managing radio resources efficiently.

To achieve the above-mentioned goal, the AN design has following attributes:

Logical isolation of access resources: AN has different types of resources. The isolation level of these resources varies according to the demands from different operators and tenants. Frequency band, time, code, equipment/site, and software are a number of dimensions, in which AN resources can be isolated. The logical isolation of resources requires resource multiplexing. The amount of allocated resources can be scaled up or down for higher utilization efficiency depending on the traffic load of each NSI.

Access differentiation: According to 3GPP's specifications for service interaction, new RAT functions consist of real time (RT) functions and non-real time (NRT) functions. In the realization of NSI in the AN, the NRT functions can apply cloudification and orchestration technologies, in order to satisfy the customization demands, such as tailored access functions and configuration differentiation. In comparison, the RT functions

should optimize the resource utilization and ensure that the security and isolation requirements of NSIs are satisfied. Specifically, the RT functions could either occupy the resources exclusively or share a portion of the resources according to the flexible configuration. Different RT functions can apply different physical layer (PHY) numerologies (e.g., frame-related parameters) to provide functional differentiation for different NSIs. The NRT functions can be deployed in a centralized unit (CU), while the RT functions are preferred to be deployed in distributed units (DUs) close to terminals. In order to deliver the fully converged 5G vision, a seamless resource management covering both fixed line and mobile access is required.

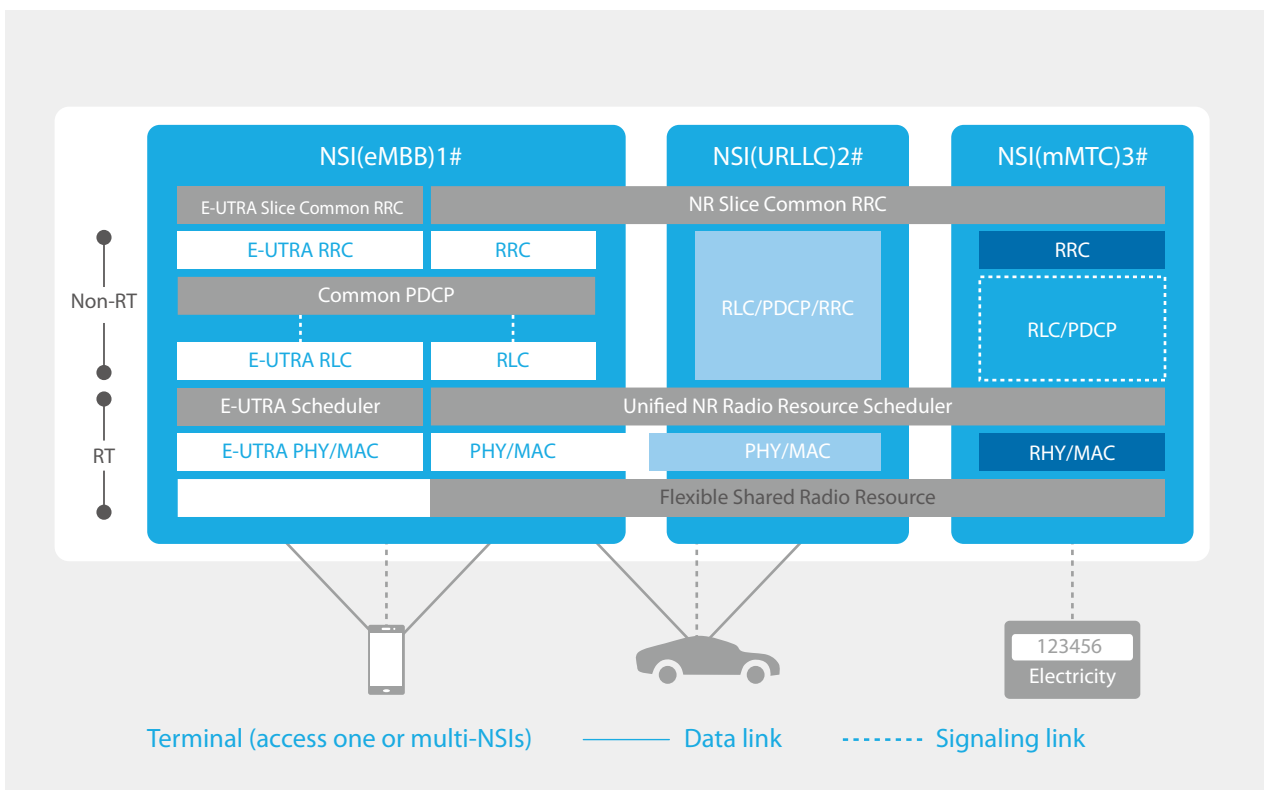
Figure 4 shows the logical architecture of network slicing support in a radio access network (RAN). The architecture has the following attributes:

Co-existence of independent and shared functions: Different NSIs can either share the same functions or have dedicated functions as shown in

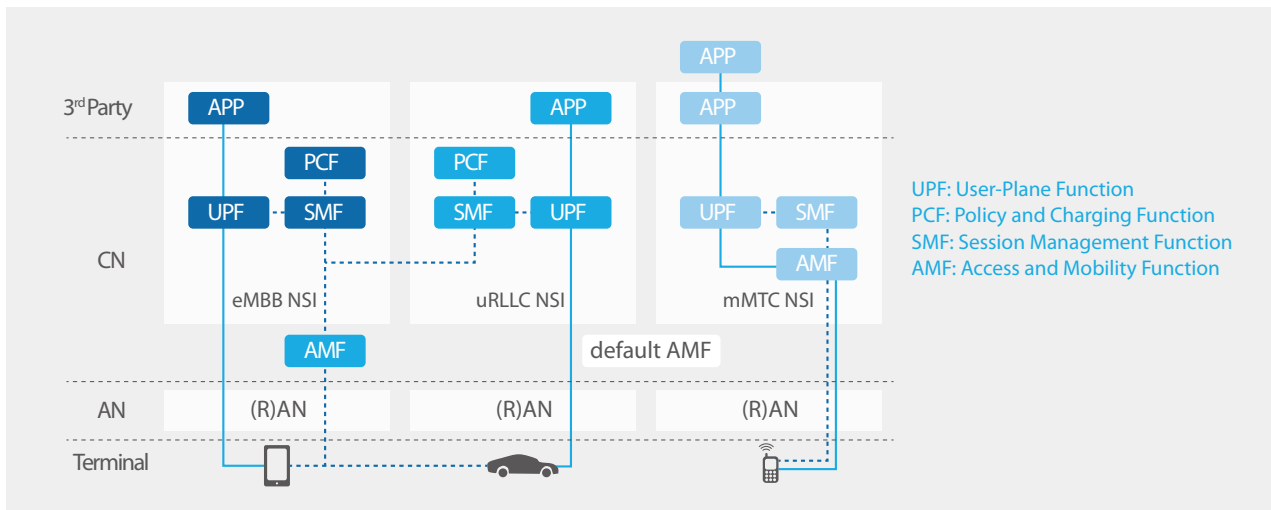
the Figure 4. New RAT features flexible air interface designs and a unified media access control (MAC) scheduling to support different network slice types. Such a combination allows time- and frequency-domain resource isolation without sacrificing resource efficiency.

Tailored functions: The protocol stack can be tailored to meet the diverse service requirements from different NSIs. For instance as shown in Figure 4, layer 3 (L3) radio resource control (RRC) functions can be customized in network slice design phase. Layer 2 (L2) can have various configurations for different NSIs to meet specific requirements for radio bearers. In addition, layer 1 (L1) uses flexible numerology to support different network slice types.

Heterogeneous access: An NSI may contain different types of ANs, such as 3GPP new RAT and Non-3GPP WLAN. Consolidating fix line and mobile access in 5G is a desirable approach, which also requires further updates on the architecture design.



· Figure 4: Radio access network (RAN) architecture with network slicing support example



· Figure 5 Core network (CN) architecture with network slicing support example

Business aspects demand flexible deployment of NSIs. Hence, NSIs may be available only in part of the network. It is possible for an access node to support one or multiple NSIs, wherein NSIs can be dynamically added or removed. Moreover, NSIs supported by an access node can significantly differ from its neighbors. In order to guarantee service continuity, terminals and access nodes should be aware of slice availability, especially in order to facilitate terminals with mobility features.

4.3.2 Core Network

The CN is considered as the most critical and essential part in network slicing. The CN provides the network services for the tenants and their end users, for instance control plane functions (mobility management, session management, policy control, and charging) and user plane functions (data forwarding). To meet the diversified demands of vertical industries, the network service customization and on-demand deployment are the key concepts that need to be reflected in the CN design. The CN is envisioned to have the following attributes:

Cloudification: Cloud native technology adoption will enable the future CN to support a large number of NSIs over common infrastructure, which is based on a three-tier DC networking mode. The bottom layer is the edge DC that is

close to or on the access sites. The second layer is the local DC, and the upper layer is the regional DC. Taking advantage of cloudification and virtualization, it is possible to let the NSIs “breathe”, in the sense of scaling the resources up and down according to the real traffic needs.

Modularization: Compared to the 4G CN, the control plane and user plane functions will be further split and partitioned into fine-grained functional modules. These modules can be customized and flexibly combined in an NSI in order to meet specific functional and/or performance requirements.

Service-Orientation: In 5G, CN will apply service-based architecture to realize flexible orchestration of network functions, services as well as capabilities, which aims to satisfy diverse requirements from different tenants with high network operation efficiency.

As illustrated in Figure 5, the CN architecture defines the services provided by different network entities and their corresponding relationships. Such services include control plane and user plane services. The CN defined for 5G should allow the coexistence of logically isolated NSIs for different tenants on a common infrastructure. The CN makes the decision of NSI selection for a terminal that attaches to the network, which could

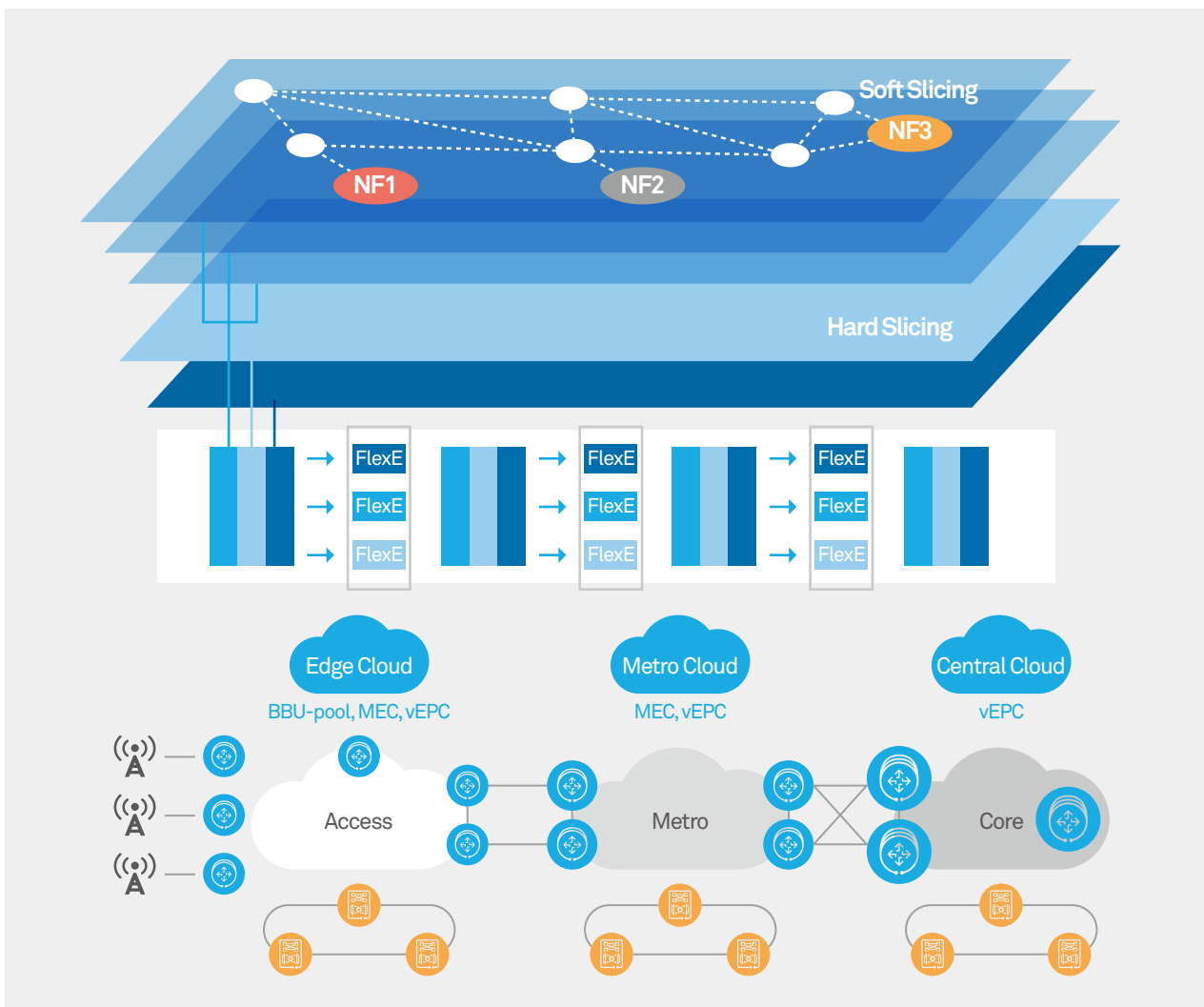
be based on the assistant from the terminal. Each NSI has its specific network topology, network functions, and allocated resources. The CN adopts a service-based architecture design methodology, which includes a unified database to enable a coherent view, and a programmable user plane to support network slicing. The CN architecture promotes the simplification of signaling interactions, enables the distribution of network functions, and allows customized network function placement (e.g., close to the AN to reduce latency). In order to consolidate wireless and wireline access using the same CN, it is also relevant to consider “access-agnostic” in the design of CN functions and architecture.

4.3.3 Transport Network

The TN comprises of different segments such as Fronthaul, Backhaul, metropolitan area network (MAN), backbone network. To support network slicing, the logical architecture of the TN is shown in Figure 6, which could coordinate protocols used by different network segments on a data path and support different types of isolations , e.g. via soft slicing method upon resource sharing or via hard slicing method upon dedicated resource.

The logical architecture shown in Figure 6 has the following attributes:

Fusion of multiple technologies: Since a single technology may not fulfill the diverse demands of



· Figure 6: Transport network (TN) architecture with network slicing support example

vertical industries, multiple technologies will be used together, for instance converged IP and optical technology in the TN domain. Each segment allows for different transmission technologies in parallel, such as optical wavelength division multiplexing (WDM), Ethernet over fiber or variants of passive optical network (PON) technologies based forwarding. Based on the features of different technologies, it is possible for the TN to provide isolation, data rate and latency differentiation for different NSIs.

Convergence of channelization and packetization: Packetization enables statistical multiplexing to increase network usage efficiency. In comparison, channelization provides necessary isolation among services. To enable multiple NSIs that carry different type of services, it is necessary to integrate these two technologies, which can divide a physical Ethernet port into multiple sub-channels for transmission, for example flexible Ethernet (Flex-E).

Enabling technologies for low latency: Network with extreme QoS demands, such as guaranteed ultra-low latency, are the key beneficiaries of network slicing. The TN requires further innovation to provide breakthroughs in latency reduction. Latency-oriented resource planning and routing, flexible Ethernet, optimized queuing in chips, and the optical network technologies are some examples. Such enabling technologies often produce high cost, which may only present on a small number of NSIs. Current research efforts are focusing on various technologies to minimize protocol stack size, to reduce transmission distance, to clear congestion, and to cut single point processing time.

Integration of TN and compute: Close integration of the compute resources from the DC with the TN will enable further optimizations and additional options for network service delivery. This also helps to eliminate various different understandings in fixed and mobile networks, thus being an important building block for FMC.

Resource isolation through virtualization and resource pooling: The TN supports network slicing based on virtualization and resource pooling. These two technologies reside in different layers of the TN. For example, virtualization technologies such as virtual private networks (VPN) and virtual local area networks (VLAN) can isolate subscriber's data transmission in the TN. A logical router can create multiple routing systems within a single router, and is another example of such network virtualization. Software-Defined Networking (SDN) technology enables operations of virtual networks on an SDN controller. As for resource pooling, WDM creates multiple optical transmission channels in the optical network. Flex-E could be one solution for pooling of link resources by partitioning a single Ethernet port into multiple sub-channels.

Customizability: The TN provides various customization methods for network slicing, in terms of logical topology, route selection, and protocol design. The TN offers NSI tenants a customizable topology with logical nodes and links. With a customizable route, different NSIs can adopt different routing policies and enable different routing protocol suites to fulfill the specific requirements of multiple services. The protocol in the TN is also customizable. For example, subscriber data can be based on different IP and Ethernet variants.

Open operation capabilities: SDN technology allows the TN to provide northbound interfaces. These interfaces permit tenants to access the logical resources in the TN. This is particularly useful for tenants to integrate their applications with networks automatically. Open operation capabilities include the processes of obtaining topologies, nodes, and links, as well as the operation of route forwarding policies, node forwarding entries, and ports.

4.3.4 Terminal

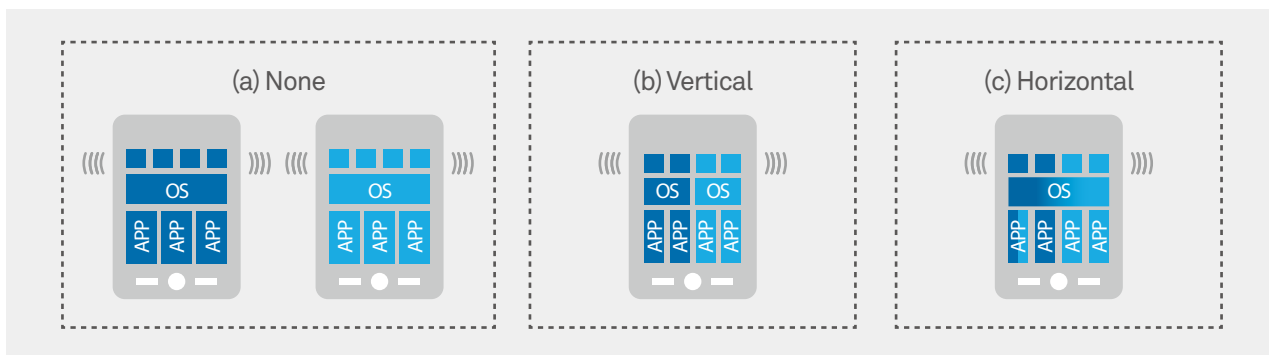
Terminal is an indispensable aspect of network slicing. 5G terminals will be diversified, from low-cost devices to high-end equipment (smart phone, car, industry robot, drone, etc.). Terminals will also have different ways to support network slicing. In general, we envision the following three scenarios shown in Figure 7.

- In the first scenario, terminals are simple devices with limited capability (e.g., sensors). This is in contrast to the second and third scenarios in which a terminal could connect to multiple NSIs simultaneously. In this scenario, network slicing related information could be pre-configured in the terminals.
- The second scenario represents a solution,

where one physical terminal contains multiple logical terminals and each logical terminal is associated with one NSI. The operating system (OS) and applications are all network slicing specific.

- For the third scenario, the OS and applications are aware of different NSIs. Terminals could configure and schedule their own resources to meet the SLA requirements of different NSIs. Hence, in this scenario, the OS is essential for supporting network slicing, especially to fulfill the various security requirements.

In addition to connecting to the network, terminals could also communicate with other terminals directly engaging in device-to-device (D2D) communications.



· Figure 7: Terminal design scenarios to support network slicing

4.4 Technology Evolution

In the early stage of network slicing deployment, there could be only a few NSIs. The deployment may occur in a semi-automatic mode. As the number of NSIs increases and scenarios, such as, dynamic instantiation of NSIs or runtime adaptation of the deployed NSI emerge, more advanced technologies will be desired to support network slicing and its further evolution. Specifically:

- Management functions will become real-time, implying that the difference between management and control will gradually disappear. Some management functions will be tightly

integrated with the NSIs as well as the network infrastructure.

- In current networks, technical domains are normally coordinated via centralized network management system. In 5G, performing real-time cross-domain coordination through distributed lower layer such as control plane would be possible, with potentially unified control logic of different domains.

- Advanced automation and AI algorithms can be applied in a unified, “holistic” network manner, which could be scalable and flexible, and which might then achieve runtime deployment and adaptation of NSIs.

5. Use Case for Service-Guaranteed Network Slicing

Automotive industry is one of the prime tenants for 5G. Its V2X services are well-accepted 5G use cases. Such services fall into three example categories:

- Entertainment: Example use case is in-car video service for the passengers.
- Driving assistant: This is a service with high requirements on latency and reliability. Example use cases are See-through, providing HD dynamic map services with environmental awareness, or real-time sharing of sensor information.
- Efficiency and comfort for fully-automated driving: This is a URLLC service. Example use cases are high-density platooning, cooperative intersection control, etc.

On this basis, the following use case is highlighted herein. An automotive provider offers in-car services cVideo (entertainment), cMap (driving assistant), and cDrive (efficiency and comfort for fully-automated driving) to its end users. The automotive provider cooperates with an operator as shown in Figure 8, who provides the network solution for the automotive provider as follows:

- Step 1: The automotive provider declares service requirements to the operator, based on which, the two parties negotiate and sign the

SLA contract via the operator's BSS. The automotive provider becomes the tenant for the operator.

- Step 2: The BSS triggers the design phase by using the Network Slice Template Designer in SSS. Based on the requirements from the tenant, the operator decides to use three NSIs, namely NSI 1, NSI 2, and NSI 3 to accommodate the cDrive, cMap, cVideo and services from the tenants, respectively. These three NSIs can have different functions, configurations, AN deployment, topology, SLA, O&M requirements, and resource usage strategies. For example, to support URLLC service, NSI 1 requires to have customized protocol stack in AN. For instance, reduction of the transmission time interval (TTI) at L1 and abandonment of a number of functions on L2 and L3 help fulfill these requirements. Part of the NSI 1's CN functions will be deployed close to the AN. In comparison, NSI 2 and NSI 3 do not pose specific demands for AN, and they can share the same AN. However, NSI 2 and NSI 3 still require separate CN that suit individual demands.

- Step 3: After the design phase, the BSS triggers the Cross Domain Slice Manager in SSS to deploy the NSIs.

- Step 4: If the tenant plans to operate services crossing multiple operators' administrative domains, this may even require cross-country

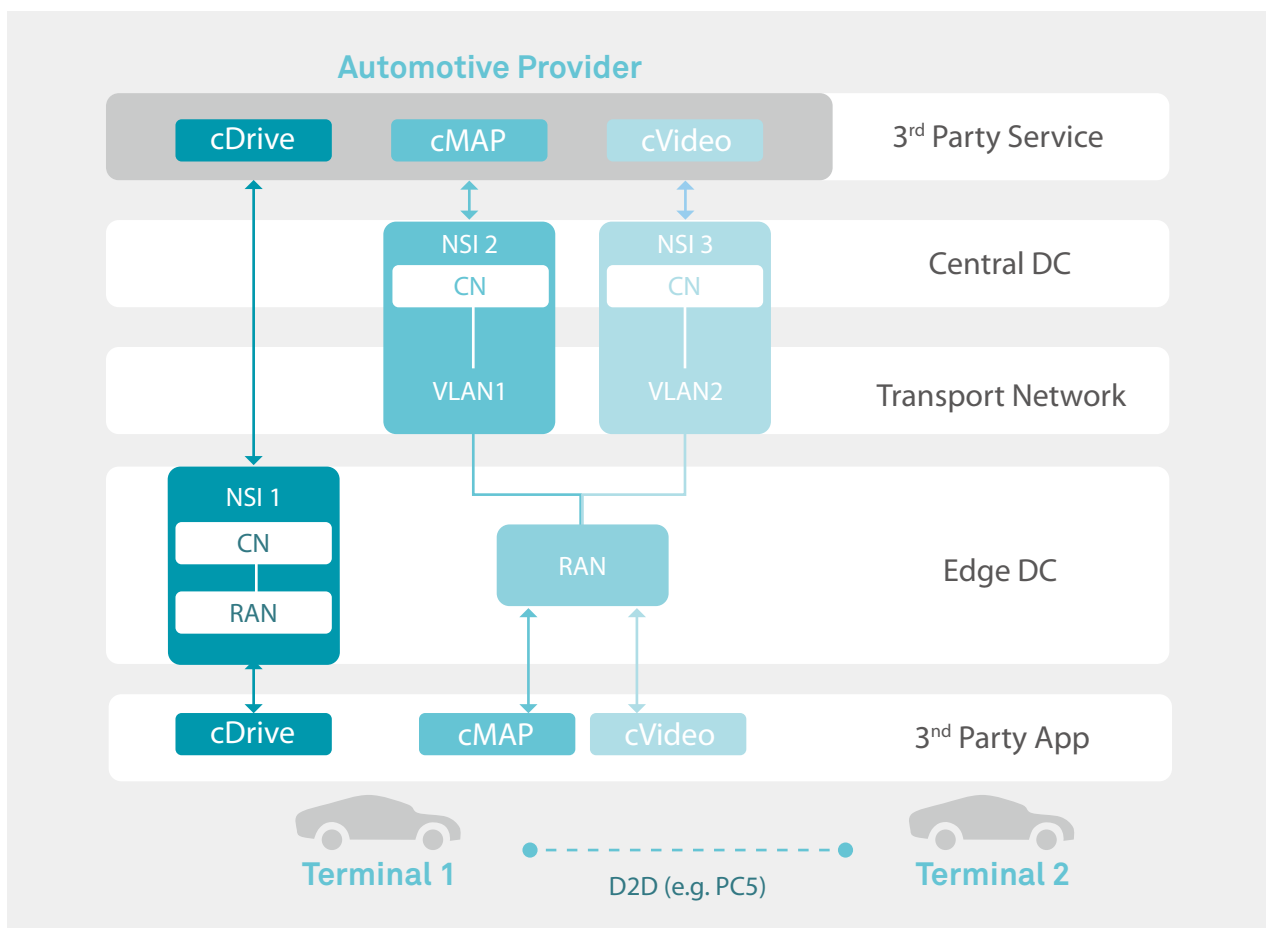
operations. The SSS will perform cross-domain collaboration with the SSSs from other administrative domains for NSI deployment.

- Step 5: If customization of TN is required, the SSS and TN-DSM will first cooperate on network planning. Multiple TN-DSMs will be involved if the equipment used in the TN is from different vendors. If no customization of the TN is required, this step will be skipped.
- Step 6: The SSS, AN-DSM, and CN-DSM jointly operate to deploy the NSIs in AN domain and CN domain, respectively. The SSS performs overall mapping and coordination among different

technical domains in order to provide guaranteed services via NSIs.

- Step 7: After the successful deployment of the NSIs, the BSS activates the usage of the NSIs.

The end users of the automotive provider can access the deployed NSIs after a successful authentication. The operator could offer the automotive provider O&M system for each NSI. The automotive provider can perform online service update and monitoring (e.g., the number of end users connected to the NSIs, alarms, geographical distribution of the end users, data traffic load, and latency).



· Figure 8: Network slice instance (NSI) deployment for V2X services

6. Summary and Suggestions

Service-guaranteed network slicing is a promising feature for 5G. It requires supports from a complex system, which calls for a plethora of enabling technologies, well-developed standards, and an increasingly mature industry chain. We anticipate tight cross-industry

collaboration among operators, vendors, and vertical industries to achieve such a challenging goal. The joint collaboration involving the following aspects will accelerate the growth of 5G industries and social economy, as well as help to create a more open and healthy ecosystem.

Establish Common Understanding among Industries

Firstly, partners from the vertical industries, service providers, operators, and vendors shall lay the cornerstones of service-guaranteed network slicing by clearly defining a set of requirements, scenarios, technologies, and strategies. The differences between the vertical industries and telecommunication industry leave a huge gap between these parties in terms of technical knowledge and business understanding. It is important to establish a common understanding, for instance via a network slicing industrial platform, among participants with different backgrounds, in order to effectively communicate and combine efforts in network slicing technology and business development.

Invigorate Technology Innovation

Secondly, network slicing requires new thinking in system architecture and protocol design, which should be flexible, efficient and future-proof. The envisioned cross-layer technologies coordinate multiple domains such as access network (AN), core network (CN), and transport network (TN).

Each of these domains comprises of functions, platforms, as well as the third-party applications. Global operators and vendors should accelerate the innovation pace in response to rising demands and challenges for network slicing.

Promote Cross-SDO Cooperation

In addition, standards developing organizations (SDOs) and open source communities shall collaborate in releasing technical standards and reference implementations for network slicing. Service-guaranteed network slicing requires cross-service, cross-layer, and cross-domain management. Currently, the standards for various domains are set by different SDOs. For example, 3GPP sets the standards for the AN, CN, security, and network management in mobile networks. IETF, ITU-T and BBF are likely to work on TN specifications. ETSI NFV provides the framework and management mechanisms for NFV, while other open source efforts like Open Platform for NFV (OPNFV) have provided NFV with implementation frameworks. The unification and collaboration of the standards in different domains are needed to facilitate the development of network slicing.

Promote Cross-Industry Collaboration

Because different vertical industries have different understanding to apply network slicing for their services, it is not feasible to bring success for all vertical industries at once. To start with, it is very important to introduce service-guaranteed network slicing to at least one vertical industry via trial or commercial deployment. Large-scale deployment is not only used for demonstration purpose, but also enables and incubates the potential network slicing innovations for vertical as well as telecommunication industries. The successful experience from one vertical industry can further attract other vertical partners, which, in turn, accelerates the full commercialization of network slicing. Cross-industry collaboration also helps foster interoperability testing and certification to advance the network slicing ecosystem.

Service-guaranteed network slicing in 5G is considered as a key driving force to realize the “Internet of Everything” vision for vertical industries. All parties shall engage in productive and beneficial pursuits, ensuring the successful implementation of network slicing and corresponding enabling technologies in the new era of 5G.

Glossary

Acronym	Definition
AI	Artificial Intelligence
AN	Access Network
AR	Augmented Reality
BSS	Business Support System
CN	Core Network
CU	Centralized Unit
DC	Data Center
DSS	Domain Slice Support System
DU	Distributed Unit
E2E	End to End
eMBB	Enhanced Mobile Broadband (a.k.a. extreme Mobile Broadband)
Flex-E	Flexible Ethernet
FMC	Fixed Mobile Convergence
HD	High Definition
IoT	Internet of Things
IoV	Internet of Vehicles
KPI	Key Performance Indicator
MAC	Media Access Control
MAN	Metropolitan Area Network
mMTC	Massive Machine Type Communications
MPLS	Multiprotocol Label Switching
NaaS	Networks as a Service
NFV	Network Functions Virtualization
NSI	Network Slice Instance
NSM	Network Slice Management
O&M	Operation and Maintenance
OS	Operating System
OSS	Operations Support System
PHY	Physical Layer
RAT	Radio Access Technology
RRC	Radio Resource Control
SDN	Software-Defined Networking
SDO	Standards Developing Organization

SLA	Service Level Agreement
SSS	Slice Support System
TN	Transport Network
TTI	Transmission Time Interval
TTM	Time To Market
URLLC	Ultra-Reliable and Low Latency Communications
V2X	Vehicle to Everything
VLAN	Virtual Local Area Network
VNF	Virtual Network Function
VPN	Virtual Private Network
VR	Virtual Reality
WDM	Wavelength Division Multiplexing

Key Contact Persons

Company Name	Contact Persons
China Mobile Communications Corporation	Wei Chen: chenweij@chinamobile.com Xiaowen Sun: sunxiaowen@chinamobile.com
Huawei Technologies Co., Ltd.	Xueli An: Xueli.An@huawei.com Shuo Wang: Saber.wangshuo@huawei.com
Deutsche Telekom	Burkhard Alfert: burkhard.alfert@telekom.de Franz Seiser: franz.seiser@telekom.de
Volkswagen	Thorsten Hehn: thorsten.hehn@volkswagen.de Teodor Buburuzan: teodor.buburuzan@volkswagen.de Steffen Schmitz: steffen.schmitz@volkswagen-infotainment.com Roman Alieiev: roman.alieiev@volkswagen.de Andreas Kwoczek: Andreas.Kwoczek@volkswagen.de

Copyright © China Mobile Communications Corporation, Huawei Technologies Co., Ltd., Deutsche Telekom AG, Volkswagen, 2017. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of China Mobile Communications Corporation, Huawei Technologies Co., Ltd., Deutsche Telekom AG, Volkswagen.

NO WARRANTY

THE CONTENTS OF THIS MANUAL ARE PROVIDED "AS IS". EXCEPT AS REQUIRED BY APPLICABLE LAWS, NO WARRANTIES OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, ARE MADE IN RELATION TO THE ACCURACY, RELIABILITY OR CONTENTS OF THIS MANUAL.

TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, IN NO CASE SHALL HUAWEI TECHNOLOGIES CO., LTD BE LIABLE FOR ANY SPECIAL, INCIDENTAL, INDIRECT, OR CONSEQUENTIAL DAMAGES, OR LOST PROFITS, BUSINESS, REVENUE, DATA, GOODWILL OR ANTICIPATED SAVINGS ARISING OUT OF OR IN CONNECTION WITH THE USE OF THIS MANUAL.

