GTI

5G Network Architecture and Capability Customization for Enterprise Network

White Paper

GTI

http://www.gtigroup.org

5G Network Architecture and Capability Customization for Enterprise Network

WHITE PAPER

V1.0

Version	V1.0
Deliverable Type	□Procedural Document √Working Document
Confidential Level	□ Open to GTI Operator Members □Open to GTI Partners √Open to Public
Program Name	GTI eNA Program
Working Group	
Project Name	
Source members	CMCC,Huawei,ZTE,Ericsson,Nokia, CATT ,Baicell ,Lenovo
Support members	
Editor	
Last Edit Date	04-22-2020
Approval Date	

Contents

1 Background

1.1 Industry-specific Private Network Requirements

It is expected that by 2021, more than 70% of enterprises will use cellular mobile networks in production environments to improve efficiency and profits. Therefore, mobile operators shall focus on gaining new profit opportunities on the 5G network infrastructure. 5G is developed for vertical industries, and private network will become an effective way to serve customers in the industries. 5G can promote the profound changes in R&D design, production control, office management and other aspects of vertical industries, and implement smart, service-based and high-end transformation of all industries. After some preliminary research on the service background, the vertical industry has different requirements for service functions, which are reflected in priority, security and mobility. Table 1.1-1 shows several typical scenarios with the following service performance requirements and indicator requirements.

Table 1.1-1 Requirements for typical vertical industry scenarios

Industry	**Industry**	**Industry**	**Energy**	**Transportation**	**Mining**	**Cloud Office**	**Healthcare**
Scenario	**Mobile robot**	**AR inspection**	**Telemetering, Telecommuni-cating and Telecontrolling over Power Distribution Network**	**Remote driving**	**Telecontrolli-ng of machines and vehicles**	**Cloud collaboration**	**Telemedicine**
Bandwidth requirements	Generally less than 500Kbps, if AGV video backhaul is required, the bandwidth is 10Mbps-50Mbps	100Mbps	Uplink: 512 kbit/s ~2048 kbit/s Downlink: 512 kbit/s	Up to over 25Mbits uplink, 1Mbits downlink	50Mbps-100 Mbps	2Mbps	≥40Mbps
Delay requirements	10 -100ms	20ms	Information acquisition < 3s, control <	<10ms	10ms-20ms	100ms	≤100ms

			1s				
Reliability	99.9% or more	99.9%	99.999%	99.999%	99.999%	Same requirements as the public network	99.999%

The above research data is taken from some scenarios described in GTI network requirement group. It can be found that the customer demand of different industries varies greatly, where they have different focus. Therefore, it is clearer and more acceptable for the industry customers for the network services that are provided in the form of private network. In order to meet the characteristics and needs of different industry applications, private networks should provide the following capabilities:

1. Customized functions, where network function and performance, access mode, service scope and deployment strategy and other functions can be customized.

2. Guaranteed quality, where the service quality requirements of the SLA can be met according to the requirements of the vertical industry.

3. It is required for agile network establishment, with rapid deployment in a short time, low cost and adjustment as needed at any time.

4. Industry-specific private network. When the private network serves a specific application scenario with isolation requirements, it can be isolated from other networks without affecting each other.

1.2 Standard Progress

Currently, in addition to the service scenario requirements to be realized with 5G private network, standardization organizations and industry alliances around the world are developing strategies to accelerate the development of 5G private network, and already have preliminary exploration results.

3GPPSA1: the working team believes that industrial scenario is an important part of 5G private networks, and the public network cannot solve the demands of factory customers. It proposes 5G private network requirements in 10 application scenarios of 5 major fields in the industrial park. For periodic and deterministic communication such as motion control, it is required that the service availability reaches three 9s to six 9s, the transmission interval is 500us to 2ms, and the end-to-end transmission delay is less than the transmission interval; for the non-periodic and deterministic communication such as emergency stop on mobile control panels, it is required that the service availability reaches six 9s, the end-to-end delay is less than 8ms, and the survival time is 16ms. For data requirements for more detailed scenarios, refer to Table A.2.1-1 of TR 22.104.

3GPP SA2: the working team designed the network architecture of private network, and proposed the concept of non-public network (NPN). NPN is a private network created through unlicensed spectrum or licensed spectrum to support typical scenarios such as enterprise communication and industrial Ethernet, and it is an important technology for mobile network to expand into the industry. The research of NPN is divided into two stages: R16 and R17. R16 defines stand-alone NPN and public network integrated NPN. For stand-alone NPN, NPN ID is assigned to the industry terminal. RAN broadcasts PLMN+NPN ID. The device makes network selection, cell selection and access control based on PLMN+NPN ID, and supports the service interoperability with PLMN. For public network integrated NPN (PNI-NPN), the concept of CAG is introduced to guarantee the exclusive access of industry customers to wireless resources. RAN broadcasts PLMN and CAG ID, the device performs network selection and cell selection based on the CAG ID, and AMF finally determines whether the device can accept the cell service based on the device's subscription information. R17 is continuing to improve the architecture and functions of NPN, including rapid subscription, third-party authentication, architecture customization, open NPN capability, independent deployment of NPN and RAN sharing of PLMN, etc.

3GPP SA3: this working team conducted security analysis and proposed security solutions for the private network architecture. In this research, we focused on how to use the existing security parameters and mechanisms to provide security protection for vertical industries, and how to provide unified security policies for the same type of devices in vertical industries, so as to avoid information leakage from weak links due to inconsistent security policies.

3GPP SA5: the working team proposed the private network management framework in R17, which is controlled by PLMN network of operators. This includes the deployment, operation and maintenance (O&M) of private networks by operators, and the provision of private network operation and O&M data to vertical industries in the form of APIs to assist the operation of private networks in vertical industries. Network domain/Cluster based NMS idea management standardization is under implementation, which will be helpful for the interfacing of industry systems and operator O&M systems.

5GACIA: the alliance is based on the industrial sector and actively inputs the identified potential 5G use cases to 3GPP, as well as the network performance requirements of each use case, to ensure that 5G gradually has the ability to serve the industry. Several white papers have been released, including 5g_for_non-public_networks_for_industrial_scenarios, which proposes four types of private network deployment:

- Independent private network, that is, NPN is a physical network with independent resources, completely isolated from the public network.

- Private network mode 1 combined with public network: Shared wireless network, that is, NPN has its own core network, only RAN and PLMN share, all other network functions remain isolated, data stream is offloaded locally.

- Private network mode 2 combined with public network: share the control planes on wireless network and core network, which are deployed by PLMN;

exclusively access to the core network user plane, which is isolated by using different APNs or slice ID.

– Private network mode 3 combined with public network: private network dominated by public network, where all networks are shared, and NPN traffic and PLMN traffic are distinguished by APN or slice ID.

1.3 Slice-based Industry-specific Private Network

Before 5G, some industries used unlicensed spectrum or industry frequency to build their own private network, but the frequency band was relatively narrow and the frequency usage was chaotic. At present, some European and American industry companies prefer to build private networks by using dedicated frequencies, but the fact is that these companies do not have professional network operation experience, and build a set of private networks for their own use, which cost is high and unnecessary. In the 5G era, operators can provide on-demand and pay-as-you-go network services through network slicing and other technologies. By using 5G's spectrum, network, operation, maintenance, operation and other resources, private networks can be implemented efficiently, at low cost and in batches. Compared with the enterprises' self-built private networks, this greatly reduces the cost of customer input and has a competitive advantage.

Both the GSMA and the China Academy of Information and Communication Technology (CAICT) have mentioned in various public forums that the fragmentation of 5G spectrum will increase the costs for the industry due to its precious spectrum resources, and the experience after the distribution of spectrum in Europe shows that it is safer and more controllable for professional operators to build the network. Therefore, operators rely on licensed spectrum and network slicing technology to build 5G private network, which is a typical development direction for enterprise private networks in the future.

Network slicing is the basic framework for serving vertical industries and the basic technology for realizing Industry-specific private networks. The slicing-based private network has the following five advantages:

– Relying on the basic conditions of the public network, it can enable rapid subscription, quick provisioning, iterative update, and provide agile and efficient advantages.

– Sharing public network resources, scale effect reduces the cost of network establishment

– With a professional O&M team, it can respond in real time, quickly solve network faults, and provide carrier-class O&M services.

– Computing power, connectivity and networking capabilities of slicing private network can be customized to enable customized and differentiated services.

– Network slice can integrate local offload capability and support edge computing

With the improvement of R15 standard, the rapid development of equipment manufacturers and the advancement of operator testing, network slicing technology have been basically enabled on the wireless, transmission, and core networks. However, the general device and industry modules have poor support for slicing, especially in the aspect of URSP (UE Route Selection Policy, the device realizes the mapping of application and slicing ID through URSP mechanism) mechanism, which has become a bottleneck restricting the promotion of network slicing technology. It is key to promoting the device industry to support network slicing in the following work.

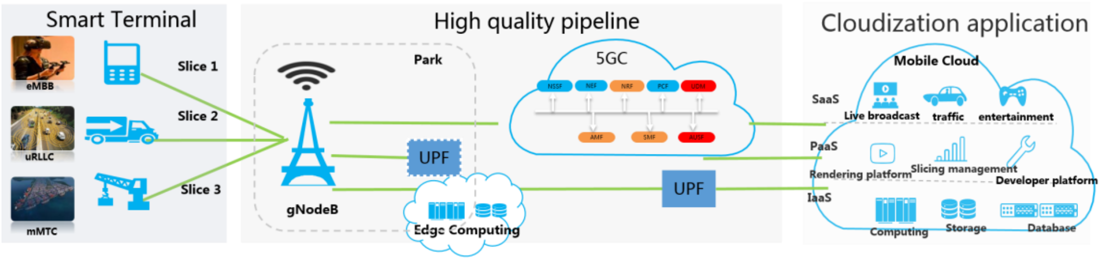

Figure 1.3-1 Build the Industry-specific private network based on network slicing

2 5G Industry-specific private network Model

2.1 5G Network Architecture

To meet the requirements of diverse scenarios and build an advanced network in the 5G era, 5G network pushes forward the transformation of architecture through the following eight technical directions.

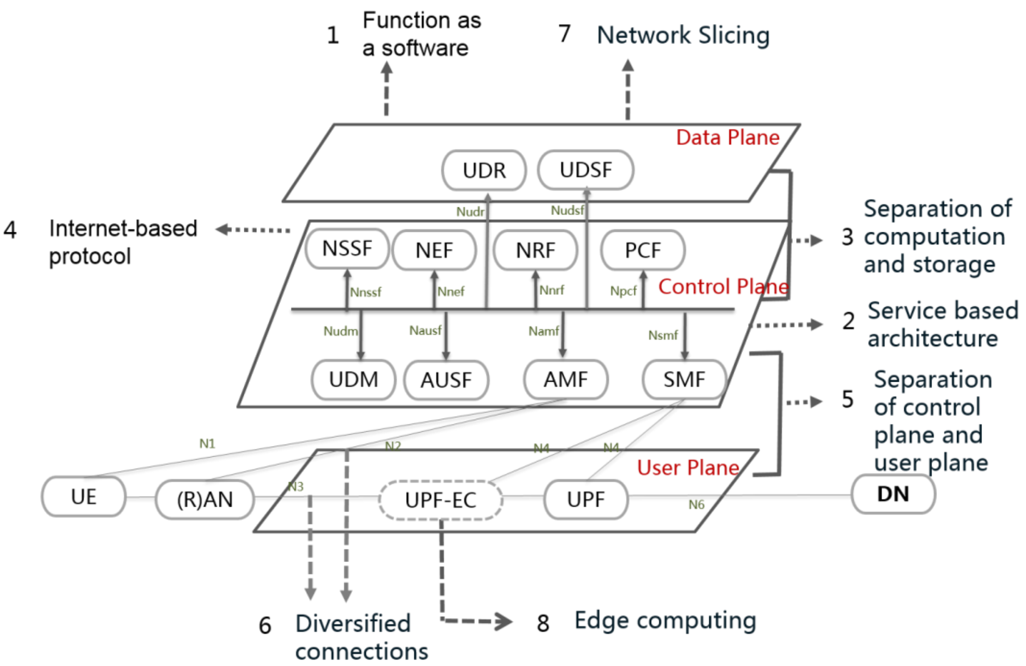

Figure 2.1-1 8 key technologies of 5G SA network

Key technology 1: function as a software. 5G enables software-defined network functions and connections through rebuilding of architecture and functions, and rebuilds 4G "network element" into 5G "network function (NF)", transforming the traditional NE-based 4G network architecture into the service-based flexible 5G network architecture.

Key technology 2: service based architecture. By referencing the mature concepts of SOA and micro services in the industry and combining with the characteristics of the telecom network, the core network is rebuilt with services to enable software-based, flexible, open and intelligent core network. 1) Network functions are decoupled into modular, self-managed “services”. The services can be independently deployed and support gray release, perform the capability customization and orchestration based on service granularity, so as to quickly and efficiently meet new service requirements. 2) API format is adopted to enable flexible invocation between services. A unified service framework is provided for service registration, service discovery, service authentication, and load management. Services can be rapidly deployed and elastically scaled in/out based on the virtualization platform.

Key technology 3: separation of computation and storage. In order to enable data sharing, big data analysis and open capability, 5G integrates all users' static subscription data. Compared with 4G, which has implemented independent data storage by functions, 5G adopts unified database and data interface to enable unified storage and access to users' static data. The user status data adopts a separate architecture: in order to enable network migration and backup disaster recovery, the context information (such as user status) in each NF of the control plane is further stripped from the NF responsible for logical processing and deployed separately.

Key technology 4: Internet-based protocol supports the new SBA architecture to have a new-generation corresponding protocol system. After the comprehensive consideration of 3GPP, the protocol combination of TCP+HTTP/2+JSON+OpenAPI was determined, so that the genes of "Internet-based" and "cloud-based" were deeply integrated into 5G mobile communication system.

Key technology 5: Separation of control plane (C) and user plane (U) functions. C/U separation is an important basis for network function optimization and flexible deployment. The control plane can be centrally deployed, controlled and optimized; the user plane functions are simplified, which can be flexibly deployed on the edge of the network to enable data forwarding with greater traffic.

Key technology 6: Diversified connections. For 4G mobility management and session management coupling, there is only one always-on session connection type; for 5G mobility management and session management decoupling, it provides service flexible networking, and establishes diversified session connections as required according to the user types.

Key technology 7: network slicing. 5G network slicing can provide customized, logically isolated and dedicated end-to-end networks including access network, transmission network and core network, which is a basic service form for vertical industries in the 5G era and meets diverse functional and performance requirements.

Key technology 8: Edge computing provides computing, storage and other infrastructure close to data sources or users, and provides cloud services and IT environment services for edge applications. Compared with the centrally deployed cloud computing service, edge computing solves the problems of long delay and large aggregated traffic, and provides better support for the services with higher real-time and bandwidth requirements. With the rapid development of 5G and industrial Internet, emerging services have a strong demand for edge computing.

2.2 Network Architecture for Private Network

Operator private network is a mobile communication service that the operator can provide for customers in dedicated industries based on authorized spectrum, with a customized range of service, customized network capability and customized isolation.

In order to achieve 5G enabling thousands of industries and meet diverse needs such as large bandwidth, low delay, high security and strong isolation, 5G SA network is divided into two relatively independent networks, C network and B network, serving public users and vertical industries respectively.

To C Network (To Customer Network), it is a network for individual users, which provides eMBB slicing service by default, interoperates with 4G, and supports roaming. To B Network (To Business Network), it is a logical private Network for ordinary industries, which can provide multiple types of network slicing services. B Network and C Network are independent of each other, and their respective control panel and user panel are independently deployed, as shown in figure 2.2-1.To support slice redirection, C network and B network share some network functions and support signaling interaction. B network is committed to expanding industry customers, and to a certain extent, it needs to be built in advance. With network slicing as a framework, multiple types of slices can be generated according to the industry requirements.

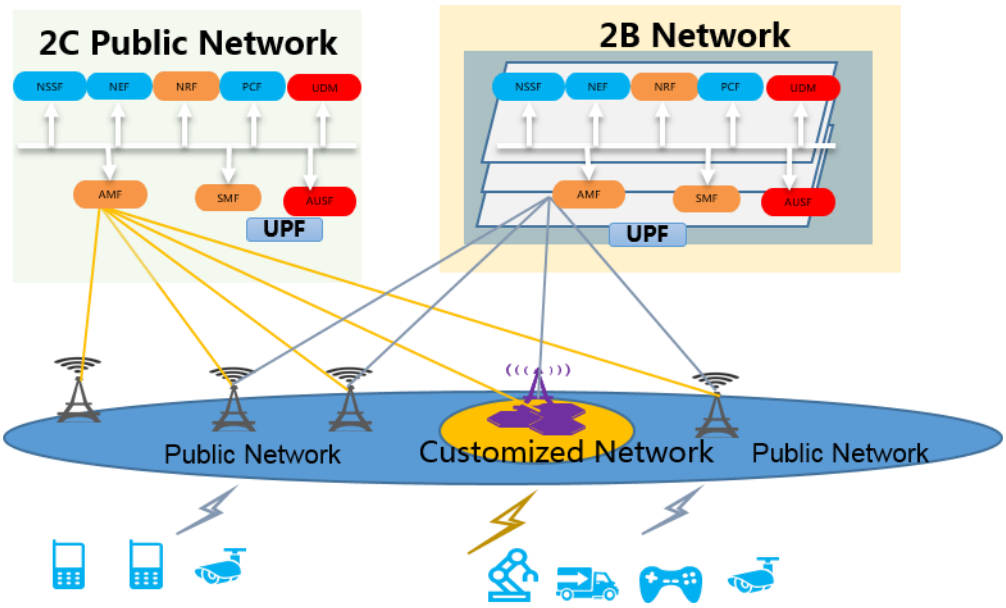

Figure 2.2-1 Relationship between 5G private network and 5G SA public network

The two networks are logically independent from each other, but are correlated. Overall networking scheme of the two networks will be described as follows from the perspective of wireless, transmission and core networks.

Wireless network: In most scenarios, the general wireless configuration can meet the service requirements of C and B networks, so it is recommended for base stations of C and B networks to use radio sharing mode, except for special industry requirements.

Transmission network: Since the resource and functional requirements of C and B networks for the transmission network can be all unified to the demand for "channels (including bandwidth, delay, jitter, security and isolation)", there is no functional difference. Currently, the transmission network can provide both soft and hard resource isolation mode, and there is no substantial gain in deploying the transmission network separately. Therefore, both C network and B network can share the transmission network, and the transmission resources can be divided as needed for the industrial application of B network through the soft and hard isolation means of the transmission network.

Core network: Due to the differences in the service characteristics, traffic model, operation management, etc. in C and B networks, core network equipment has great differences in functional requirements, equipment upgrade cycle, network data configuration, equipment O&M mode, etc., so the core networks of B and C networks are independently established, and B and C networks are interoperable to meet the needs of slice redirection.

When B network faces industrial customers, it can provide different levels of service modes according to customer needs. For example, high bandwidth and isolation requirements can be enabled by QoS, DNN priority scheduling or network slicing, or the data offload capacity of edge computing can be superimposed on the above basis to meet the data requirements of not out of the park and low delay. The wireless side can also enable higher security, high isolation, and customized services by sharing public network frequencies or using dedicated frequencies/base stations.

3 Typical Network Capabilities

3.1 Typical capabilities of wireless network

This section describes the key capabilities that wireless networks can provide for vertical industry scenarios. Industry customers can customize and acquire the types of enterprise private network services according to their unique business requirements.

3.1.1 Wireless network customization service

In terms of the radio access network (RAN), which is the terminal end of the entire cellular wireless network, not only the requirements of the business for customized private network must be met, but also RAN's own characteristics must be taken into account. In terms of both full utilization of spectrum resources, and network construction and operational efficiency, the network requirements involve a large number of logical private network scenarios.

Considering that industries pose various requirements for cellular wireless networking, there can be large-scale continuous coverage, linear coverage according to designated paths, and discontinuous coverage in multiple areas. RAN needs to form the ability to build logical networks based on logical network elements to reduce the coupling between logical private networks and physical facilities. The ability to build logical networks therefore becomes the basis for RAN to further meet various requirements such as enterprise private network service distribution, management and maintenance.

In scenarios where an enterprise terminal needs to move in a continuous wide coverage network, a logical network can be constructed according to different service types of the enterprise. This is the multi-service logic private network. For a local coverage scenario that moves only within a specified area, logic can be constructed according to the geographic scope, i.e. a regional logical private network. Alternatively, the logical network is further divided by tenants within the same geographical area, which is the tenant logical private network. These types of logical private networks may concurrently exist.

The RAN logical private network can be used as an access instance of end-to-end network slicing, or as an access part of an NPN network, or further, as an instance of RAN sharing in the case of public and SNPN overlay networking.

The isolation capabilities of RAN include operation and maintenance isolation, security level isolation, reliability and fault isolation. The isolation modes are divided into three categories. The first category uses dedicated equipment and dedicated frequencies for the most thorough isolation; the second category involves sharing of base stations and using dedicated frequency to isolate according to cell granularity; and the third category uses QOS or slice to isolate services within the cell.

The sharing capability of RAN is reflected in different types of resources, including wireless air interface and antenna resources, communication protocol layer processing resources in the device, transmission switching channels and port resources, etc.

The RAN isolation and sharing strategy needs to be established according to the requirements of the business. For example, when a business has a very high requirement for reliability guarantee, sharing of resources would affect that reliability. Some industries have their own management requirements or technical specifications (e.g. enterprise security specifications etc.) that require the processing resources isolation of communication equipment, while others do not.

For wireless spectrum resources, regardless of the isolation technology in use, the secure transmission capability is equivalent. This is because the bit stream of the cellular network is transmitted in the air, and regardless of whether the wireless resource blocks carrying the bit stream are continuous, the adjacent blocks will not affect each other. Therefore, a key determinant in the isolation strategy of wireless air interface resources is the business's requirement for service reliability guarantee.

3.1.2 Value-adding wireless network services

Compared to the public network, importance of RAN's autonomy capability is highlighted in enterprise private network scenarios. Both the priority and exclusive models are deployed in a highly decentralized manner. In the case of centralized operation and maintenance of multiple enterprise private networks, operators must consider reducing the OPEX cost of communication equipment maintenance.

Within the scope of the designated local area network, RAN-related service activities should aim for local autonomy and fast closed loop as much as possible, including the planning, configuration preparation, network optimization and maintenance of infrastructure resources and parameters. Through local autonomy of the local area networks, the layered operation and maintenance of the enterprise's private network can be enabled, so that part of the network maintenance tasks can be automatically closed and quickly resolved locally within the enterprise, thereby improving the overall management efficiency of the private network.

3.2 Transmission network capabilities

The 5G backhaul network SPN enables the transmission of related traffic between the CU and the core network, and between the CU and CU. It consists of three layers: access, aggregation, and core. Considering that the mobile core network will be developed from 4G evolved packet core (EPC) network to 5G new core network and mobile edge computing (MEC) etc., and that the core network will be cloud-based and deployed in large provincial backbone and metropolitan core data centers, while MEC will be deployed at edge data centers in metropolitan or lower areas at the same time, the metropolitan core convergence network will evolve into a unified bearer network for 5G backhaul and data center interconnection. Additionally, the transmission network can provide the survivability mechanisms such as protection and recovery according to the actual business needs, including the optical layers, L1, L2, and L3 etc., to support the high reliability requirements of 5G services.

For different applications (eMBB, URLLC, mMTC) provided by 5G, network slicing allows the transmission network to be concentrated in a unified physical network, integrating related service functions and network resources to form a complete, autonomous and independent operation and maintenance virtual network (VN) to meet the requirements of specific users and business. The key technologies for building a virtual network include SDN/NFV management and control functions and forwarding plane network slicing technology. SDN/NFV is responsible for the virtual abstraction of resources, and the network slicing of the forwarding plane is responsible for the isolation and allocation of resources to meet differentiated virtual network requirements.

The 5G transmission network SPN provides a hierarchical network slicing solution that supports hard and soft isolation to meet the requirements of different levels of 5G network slicing. For example, services such as URLLC, financial government-enterprise dedicated lines and similar services require exclusive resources, low delay and high reliability. The transmission network can provide L1 TDM isolated network hard slices; eMBB Internet access and AR/VR video services characterized by high bandwidth, delay insensitivity and dynamic burst, etc. The transmission network can provide network soft slice based on L2 or L3 logic isolation.

3.3 Typical core network capabilities

3.3.1 Network slicing

The 4G network serves mainly the general public users. The main devices connected to the network are mobile phones, unnecessary to divide the network into slices for different scenarios. The 5G network, on the other hand, is oriented to the interconnection of all things. Different terminals need to be applied in different fields, and the network performance requirements in the scenarios differ too, thus network slicing came into play.

Network slicing is an end-to-end logical private network that provides specific network capabilities. An instance of network slicing is a collection of network functions and required physical/virtual resources, which may specifically include an access network, a transport bearer network, and a core network. Network slicing can be built based on traditional proprietary hardware, or the NFV/SDN general infrastructure. It is recommended to build on a unified platform to enable low cost and efficient operation. The capabilities of network slicing can be customized. So can network performance, access methods, service scope and deployment strategy. Slicing aims to serve a specific application scenario. Different slices are well isolated from each other without interference. The SLA service quality requirements can be met according to requirements of the vertical industries.

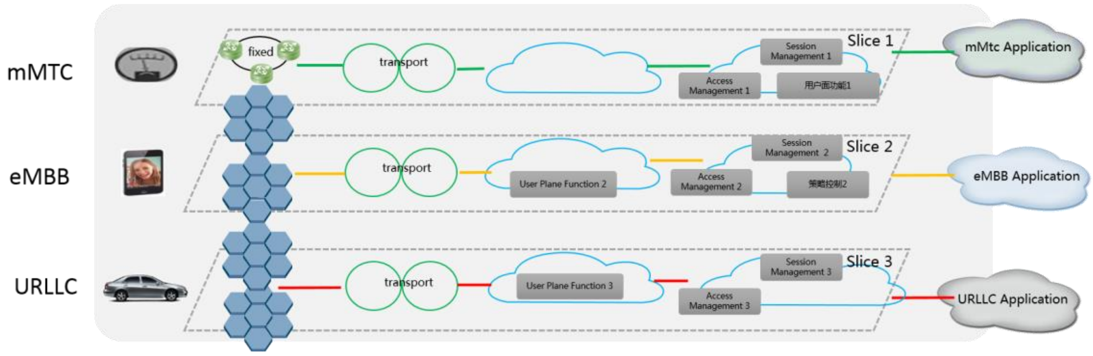

Figure 3.3.1-1 Overall concept of network slicing

Network slicing is the basic form for vertical industries in the future. When network slicing is offered as a service, industry customers can order 5G network slicing from operators based on service type, service scope, performance index requirements, and special functional characteristics. Network slicing provides network connection services, customization and value-added services. Through network slicing technology, an end-to-end service system for the entire industry and multiple scenarios can be built to enable the specialization of the public network. For operators, network slicing can help integrate the underlying network resources to enable efficient and centralized operation, maintenance and management. For vertical industries, the customized end-to-end network connection solution can well meet the requirements of various scenarios and resolve the industry's communication pain points and security concerns in precision.

3.3.2 Local offload

Right from the beginning of its design, the 5G network has supported edge computing. Based on the C/U separation architecture of the 5G core network, the user plane function (UPF) needs to sink to the edge of the network to reduce transmission delay and enable local offload of data traffic. There are three local offload methods: uplink offload, IP source address selection, and local data network. The uplink offload, i.e. ULCL of the UPF, performs offload by checking the destination IP address of the data packet according to the filtering rules delivered by the SMF. IP source address selection, i.e. BP of the UPF, offloads by checking the source IP address of the packet according to the filtering rules issued by SMF. And with local data network method, the terminal judges its own location; if the terminal is in the LADN service area, it initiates a session establishment request carrying the LADN DNN.

The 5G network also introduces three service and session continuity modes to support edge computing (SSC mode 1/2/3) to ensure user experience in terminal mobility scenarios. 5G network exposure function (NEF) supports exposure of network information to edge applications. APIs such as wireless network information services, location services, and QoS services have been defined in the edge computing system, and the wireless network information, user location information, and QoS information acquired from the mobile network are encapsulated and exposed to edge computing applications. For further information on the edge computing platform, refer to Chapter 4.

3.3.3 5G LAN

From 2G to 4G, the mobile network has provided unified access and consistent terminal management. In the advent of the 5G era, industry customers want 5G networks to provide large connections, high bandwidth, and low delay, as well as self-built local area networks to enable flexible management of terminals. For example, industry customers want to be able to specify the IP address of the terminal, require the terminal to only communicate with a specific terminal, authorize the terminal to be grouped to a specific group and dynamically join and delete from the group. 5G LAN technology is just to meet the demands of industry customers.

For the first time, 5G LAN technology introduces the concept of terminal group management in mobile networks, to support direct communication of terminals in the group. The technology is characterized by the following three features.

Feature 1: Dynamic group management. UDM maintains all contract information of terminals in a group. Industry customers can dynamically add or delete a terminal from the group through the capability exposure interface.

Feature 2: Specify the IP address of the terminal. Industry customers can configure the specified IP address segment into UDM through the capability exposure interface. When the terminal uses 5G LAN DNN to create a session, the 5G network will use the static IP address allocation method to assign the specified IP address stored in the UDM to terminals in the group. This enables the effect of industry terminals using industry intranet IP addresses, so that terminals accessed through 5G can seamlessly communicate with application servers or terminals in the industry intranet to build industry-specific wide-area "local area networks."

Feature 3: Direct communication via UPF. The data sent by the terminals in a group can be directly forwarded through the 5G gateway UPF without passing through the N6 port. When the terminals in a group access multiple UPFs, these UPFs will establish a direct connection tunnel to support direct communication. This design empowers 5G UPF with the same direct forwarding function as a router, thus no longer requiring the intervention of a router or application server on the N6 port, and shortening the data forwarding path.

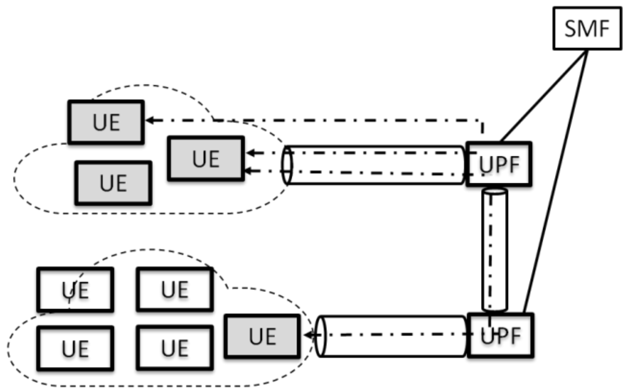

Figure 3.3.3-1 5G LAN supports direct communication between terminals through UPF

The technology has broad application prospects. For example, in the case of massive IoT terminal access, 5G LAN technology can easily locate the terminal group of customers in a certain industry, and support customers to configure the terminal independently. In industrial scenarios, by constructing different 5G LAN groups, different levels of terminals can be securely isolated.

In April 2019, China Mobile, Huawei, and Baidu successfully verified the capabilities of 5G LAN for the first time, by enabling dynamic IP address configuration of cameras in a group through Baidu Cloud, and preliminarily verified the ability of terminal group management. At the Baidu Cloud Summit later on, China Mobile, joining hands with Huawei and Baidu, demonstrated for the first time an 8K live video based on 5G LAN technology. The 8K camera was accessed through the 5G network and communicated with the Baidu cloud server. The camera and the server are both in the same local area network.

The 5G LAN technology has no impact on terminals and base stations, and only requires the core network to be modified and upgraded. It can be introduced first as a 5G network enhancement technology in vertical industry networks.

3.3.4 URLLC

URLLC (Ultra-Reliable and Low-Delay Communication) is one of the three major 5G application scenarios defined by 3GPP. Typical scenarios are:

- Industrial control: delay of 5 – 20ms, reliability of 99.999% or higher
- Smart grid differential protection: delay of 15ms, reliability of 99.999%
- Remote operation: delay of ≤20ms, reliability of 99.999%
- Autonomous vehicle: delay of 3 – 100ms, reliability of 99.999% or higher

The URLLC network is required to provide users with millisecond-level end-to-end delay and service reliability of close to 99.999%. To meet this requirement, 5G networks guarantee URLLC service experience by building redundant links and end-to-end delay monitoring.

1) Build redundant transmission links to support highly reliable data transmission

In order to meet the reliability requirements of URLLC, 5G networks support the construction of end-to-end redundant transmission links or N3 tunnel redundant transmission links.

To minimize the impact of wireless channel fluctuations, the 5G network improves overall reliability by establishing two end-to-end sessions. As shown in Figure 3.3.4-1, the end-to-end session is initiated by the UE, and two sessions are established to connect to the same data network (DN) through two NG-RANs. The transmission paths of the two sessions do not intersect. By this mechanism, both the terminal and the base station need to support dual connectivity.

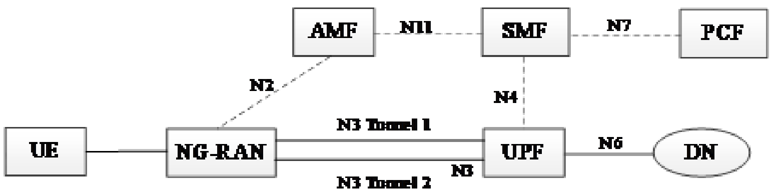

Figure 3.3.4-1 Construct redundant transmission links to support highly reliable data transmission

The 5G network also introduces the N3 redundant tunnel mechanism to enhance the reliability of the transmission link. The two N3 tunnels are established between the same RAN and UPF; but the underlying transmission links do not intersect. In order to implement this solution, NG-RAN and UPF need to construct two N3 tunnels when establishing a PDU session according to the control plane requirements. NG-RAN and UPF need to support the duplication and deduplication of uplink and downlink data packets.

The access network and core network of the 5G network can be flexibly configured to enable the service requirement of ultra-low delay. The access network can enable a one-way 0.5-1ms air interface delay by configuring an ultra-short frame structure. The core network uses UPF sinking to make UPF deployed closer to users, thereby reducing end-to-end delay. The 5G network also supports an end-to-end delay measurement mechanism. According to the requirements of third parties, real-time delay measurement of the service is performed per flow and reported in real-time. The 5G control plane can therefore be updated of the current end-to-end delay situation in a timely manner. When delay fails to meet the requirement, assist with network topology adjustment (for example, UPF reselection), or QoS policy adjustment to meet the low-delay service requirement. Delay monitoring can be performed based on QoS flow granularity or node granularity. Through redundant transmission and periodic data packet measurement, the 5G network can substantially meet the service requirements of URLLC. However, introduction of the technology requires the terminal and application layer to be modified according to the redundant transmission mechanism. The terminal and the base station need to support dual connectivity. The base station needs to be configured with a flexible frame structure to enable a 1ms unidirectional air interface delay. Periodic packet detection will also consume a lot of the network resources. But undeniably, the ultra-low delay and ultra-high reliability features brought by the URLLC technology will bring revolutionary changes to scenarios such as remote surgery and industrial control, and it will become a true high-value technology for 5G to serve vertical industries.

3.3.5 5GS supports TSN

Compared with the guarantee of URLLC technology in terms of reliability and delay, TSN technology further enhances 5G networks in terms of delay jitter and time synchronization. 5G deterministic network refers to a mobile network that provides deterministic transmission for specific applications, i.e. bounded delay, low jitter, extremely high reliability, and end-to-end high-precision time synchronization.

The 5G system has been expanded to support protocols such as IEEE 802.1AS clock synchronization mechanism, 802.1Qbv threshold control mechanism, and 802.1QccTSN configuration mechanism at the R16 stage, thereby building an end-to-end time-sensitive network.

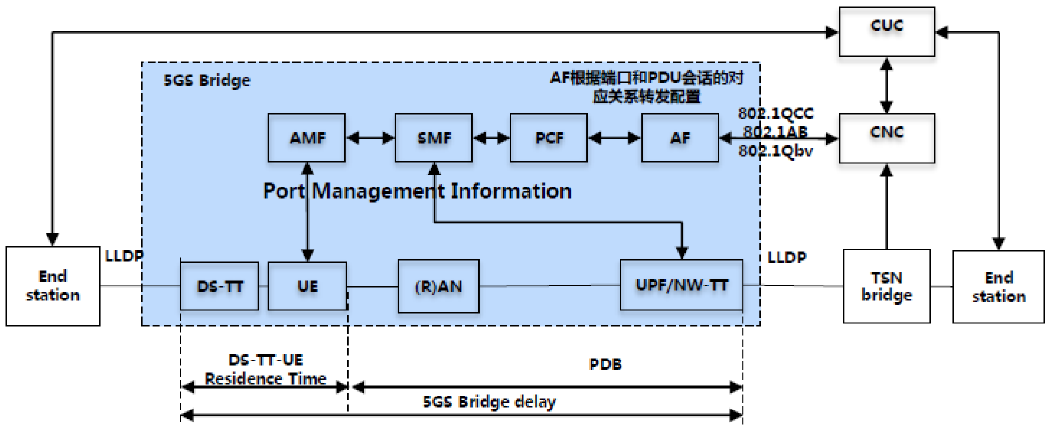

Figure 3.3.5-1 TSN flow scheduling and transmission on 5GS Bridge

Architecture enhancement: The 5G system is integrated into the TSN system as a TSN bridge. This "logical" TSN bridge includes a TSN converter for user plane interaction between the TSN system and the 5G system. The 5GS TSN converter function consists of a terminal-side TSN converter (DS-TT) and a network-side TSN converter (NW-TT).

Time synchronization: In order to enable the TSN synchronization mechanism, the entire end-to-end 5G system can be regarded as an IEEE 802.1AS time perception system. In the 5G system as a time perception system adapted to the IEEE 802.1AS protocol, there are two time synchronization domains, namely the 5G time domain and the TSN time domain. Within the 5G system, 5G GM (5G internal master clock) synchronizes with UE, gNB, UPF, NW-TT and DS-TT; only the TSN converter (TT) at the edge of the 5G system needs to support

IEEE 802.1AS-related functions such as (g)PTP, timestamp, Best Master Clock Algorithm (BMCA) etc.

Time-sensitive communication (TSC) QoS control: First, the 5G system and TSN system negotiate the QoS requirements of each TSN service flow, such as bandwidth requirements and delay requirements. The DS-TT and NW-TT in the 5G system support the store-and-forward mechanism of 802.1Qbv. Within the 5G system, the TSN service flow will be guaranteed by delay-sensitive GBR. In each cycle, the TSC QoS flow is required to transmit only one burst data block, and the size of this burst data block is used to set the Maximum Data Burst Volume (MDBV) in the QoS parameters. The Packet Delay Budget (PDB) in the QoS parameters comes from the TSN service requirements, and the PCF will decompose the PDB into CN-PDB and AN-PDB according to the network delay measurement results, where AN-PDB is used to guide the wireless processing delay. Guaranteed by the QoS GBR mechanism, despite 5G system's transparency to the TSN system, it is capable of completing the data flow scheduling transmission required by the TSN system on time.

The adaptation and interfacing work with the IEEE802.1 mainstream TSN protocol family in 5G system have been basically completed at the R16 stage, and deterministic data transmission at the MAC layer is enabled. However, due to the complexity and diversity of the IEEE TSN protocol itself, some issues arose in 3GPP standardization of the 5G TSN technology, and the work of the 5G TSN standard is advancing slowly. Consequently, 3GPP plans to launch a deterministic mechanism specific to mobile networks at R17.

5G TSN technology requires transformation of terminals, base stations, transmission and core networks. Terminals and UPFs need to support TT (TSN Translator) functions, and core network AF needs to support interfacing with the TSN system control plane (CNC/CUC) and complete protocol interpretation and parameter mapping. Furthermore, the 5G core network needs to be synchronized with the 5G master clock of the transmission network and the base station. Although it is costly to introduce TSN in 5G networks, it is foreseeable that 5G TSN technology will be widely used in industrial control, machine manufacturing, high-definition audio and video transmission and other areas. A 5G network with TSN attributes will become a truly deterministic 5G network with deterministic delay, low jitter, and high reliability, providing true SLA protection for vertical industries.

3.4 Security requirements of private networks

Vertical industries have diverse requirements for private network security, mainly due to the diversity of service requirements, the comprehensive cloudification and opening of network architectures, and the complexity of enterprises' local ICT networks. When the 5G private network enters the vertical industry, the security requirements of customers will be fully considered to provide diversified and customizable security capabilities.

Under the slice-based industry-specific private network, on the one hand, the 5G system applies more comprehensive and stricter security protection measures for the 5G network, and on the other hand, it provides the security capabilities of the slice itself for access, isolation, and management. In addition, the 5G network can combine the security capabilities provided by the operator with the network exposure capabilities, so as to efficiently, quickly, and flexibly provide the security capabilities needed by businesses in the vertical industry, and reduce the implementation costs and application development cycle of business security strategies in the vertical industry.

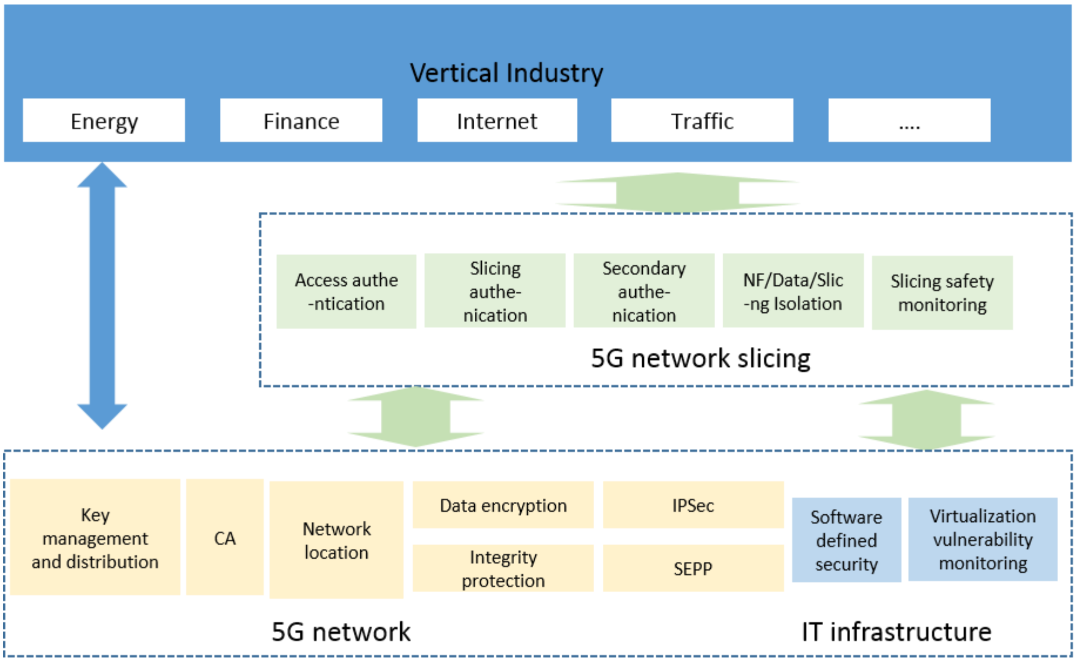

Figure 3.4-1 5G slice security capabilities for vertical industries

The security capabilities currently provided by the 5G system for private networks are as follows:

1. It provides slice-level security protection measures, including the internal security of the slice, as well as access security, transmission security, and management security:

   a) It provides slice-level user access security. Based on the 3GPP unified authentication framework, it can perform two-way authentication of terminal identities in various application scenarios to ensure the legality of access terminals and access service networks. 5G provides users with access slice authentication and privacy protection for slice selection auxiliary information to ensure that legal user terminals access legal slices while shielding the relevant information from non-target terminals and network devices. In addition, 5G networks can also provide data of air interface transmission with security encryption and integrity protection capabilities, supporting the security capabilities of various types of data transmitted by industry users on 5G networks to prevent eavesdropping and tampering;

   b) It implements the isolation of slices and establishes a three-level three-dimensional security isolation system. By implementing isolation between slices, isolation between slice networks and users, and isolation between network elements within slices, network-specific, customized, and exclusive connections are enabled. On this basis, according to the production control of the vertical industry and the bandwidth requirements of the collected information, slices with different attributes are divided to enable the logical or physical isolation between the slices, so that the private network is dedicated for its use. Sensitive network elements can be deployed in physical locations and equipment of high security level as required. Further, according to the requirements of vertical industries, the bandwidth and frequency band of the base station are set, an isolated 5G private network is constructed, and terminals not in this industry are denied access to the 5G private network to ensure wireless side security capabilities;

c) For slice management, real-time slice security monitoring capabilities, emergency response and fault recovery capabilities are provided to detect security incidents when the enterprise network is attacked or illegally invaded, switch service nodes, and recover the network elements that have been attacked and failed, thereby ensuring system availability. Additionally, 5G network provides security protection for slice management, including two-way authentication, authorization, and log auditing of slice management operations, encryption, integrity protection, anti-replay protection of slice management operation command transmission, check and verification of slice configuration, and data clearance after the slice is terminated.

2. It provides slice-level security service capabilities, combining security capabilities that operators can provide (e.g. authentication, authorization, auditing, intrusion detection, dynamic security protection) and software-defined network (SDN), service function chaining (SFC), SDSec (Software-defined Security) and other technologies to provide security services to vertical industries. In network slicing for vertical industries, three methods are adopted to provide security capabilities:

a) When signing a slicing service with a vertical industry, the security configuration parameters (e.g. encryption algorithm, key length, user plane integrity protection, etc.) for a specific slice are determined in the form of SLA, as well as the security baseline. When necessary, vertical industries can apply to operators to adjust the security configuration and parameters;

b) The operator's network and security capabilities (e.g. cellular-based positioning, identity management, session key negotiation, mobility management, and restrictions etc.) are exposed by the network exposure function (NEF) with a programmable interface API (Application Programming Interface) to application developers in vertical industries; and application developers can call it on demand in service logic.

c) Security devices or security function modules on the network side are deployed to allow specific application traffic pass through these security modules in turn through flow scheduling, thereby building defense in breadth and in depth. For example, the operator network can provide anti-DDoS attack, firewall-based access control, IPS, and WAF functions for inbound traffic of the slicing application.

Furthermore, where MEC servers are introduced into the network, in addition to isolation methods such as virtual machines and containers, IPSec, TLS and other encrypted transmission methods, firewalls, layer 2 and layer 3 data isolation and other logical isolation general security measures can prevent attacks on the MEC servers, mobile systems and third-party servers, while establishing a unified and reliable security evaluation system to assess the security of applications installed in the MEC system and the security between application interfaces and the network.

4 Platform Empowered Industry Applications

Multi-access Edge Computing (MEC) migrates computing storage capabilities and business service capabilities to the network edge, enabling applications, services, and content to be deployed in a localized, close-range and distributed manner to meet certain service requirements for technologies and scenarios such as eMBB, URLLC, and mMTC in the 5G systems. Meanwhile, by using MEC, network data and information can be explored to the fullest extent to enable perception and analysis of network context information, and potentially open to third-party service applications, thereby effectively improving the level of network intelligence and promoting the deep integration of network and service.

The edge computing platform provides a deployment and operating environment for the localization of various 5G new service applications (e.g. AR/VR, local campus applications, etc.). Therefore, edge computing platforms play a particularly important role in vertical industries. Details are as follows:

- Reducing delay

For URLLC services, the service's delay requirements can be very stringent. The UPF and MEC platforms can sink closer to the network edge according to their delay requirements, thereby minimizing the impact of transmission delay.

- Providing abundant storage resources and improving network utilization

In live games, edge computing can provide CDN with rich storage resources as well as audio and video rendering capabilities closer to users, making new service models such as cloud desktops and cloud games possible. Notably, in AR/VR scenarios, the introduction of edge computing can greatly reduce the complexity of AR/VR terminal devices, thereby reducing costs and promoting the rapid development of the industry as a whole.

- Interconnecting with 5G network

NEF can be deployed on the MEC platform to enable local exposure of 5G network information. On the one hand, NEF transmits the relevant measurement information of the UE and the service flow, e.g. the real-time location of the UE and the quality of the wireless link, to the MEC server, in order to intelligently analyze and abstract this measurement information and optimize service performance. On the other hand, NEF transmits application service-related information such as service duration and service cycle perceived by MEC to the network, for the latter to perform perceptual analysis on the information and further optimize UE resource configuration (for example, allocate appropriate bandwidth resources to VIP users) and session management.

- Open edge computing

The edge computing platform is an open platform, which is reflected in the openness of edge network capabilities, platform management, and platform services. Building open edge and incubating innovative services with partners are the key enablers of operators.

- Enhancing local data processing capability

In intelligent manufacturing, factories use edge computing platforms for local data collection and real-time processing such as data filtering and cleaning. At the same time, edge computing can support the unified access of fragmented industrial networks. Some factories are still trying to use virtualized technology software to implement industrial controllers to perform centralized and coordinated control of production line robotic arms.

- Improving privacy and security

The MEC platform is placed near data sources (IoT equipment, workers, operation and maintenance personnel etc.) and consumption points, and processes and feeds back on collected data locally in real time without going through the traditional core network, thus satisfy the reliability requirements and data privacy requirements.

5 Considerations of Operation and Operational Maintenance

5.1 Flexible relationship between operation and operational maintenance

Operation: The operation of the private network for enterprises requires greater flexibility in comparison with the public network. Private network needs to adapt to specific industry scenarios in terms of scope and attribution, and several typical operation models are derived based on industry categories. From the architectural perspective, it is recommended to consider the operation issues in several levels, among which the coupling with the network communication requires unified deployment of the operator. Anything above this level is determined between the industry and the operator through negotiations, on an overall premise that the operator has complete operational capabilities, and the ability to integrate other operational software. Network communication capabilities are an integral part of business requirements. Enterprises need to consider not only direct communication, but also integration and interfacing, evolution, and migration issues. Also, it is very costly for enterprises to switch their fundamental service providers. With the continuous development of business, its ICT demands will gradually evolve. Communication is never an isolated requirement for enterprise customers.

Operational maintenance: At the foremost is the maintenance of communication functions and network resource entities, among which the network resource nodes, site addresses and physical entities are all key basic capabilities. Based on these capabilities, others capabilities required by the industry such as positioning can be derived. In logical private network scenarios in particular, operational maintenance involves two functions; namely, the maintenance and optimization of communication service quality interpreted by key indicators, and the centralized management and control of network resources corresponding to other industry capabilities.

The relationship between operation and operational maintenance: The relationship between enterprise network operation and network operational maintenance needs to be more flexible and able to gradually expand in line with business changes. On the one hand, exploring new operation models in the enterprise market requires innovative charging models and service packages, as well as timely adjustments and changes in the functions of entities related to network operations, and these changes should not result in unpredictable effects on network quality. Consequently, this requires the decoupling operation and operational maintenance. On the other hand, the enhancement of some operational maintenance functions can better support operation expansion, including operational maintenance capabilities of visualization and automation, which help to integrate the capabilities of cellular networks into the digital service flow of enterprises, such as a highly automated network helps operators establish a centralized operating platform.

Enterprises themselves are experiencing transition from enterprise private networks to broadband and converged service needs, accepting operators to enter the enterprise service market. On the one hand, network operation and operational maintenance are the core capabilities of operators that provide confidence for enterprises in the reliability, stability and security required for normal operation. On the other hand, enterprises can also rely on operators to gather forces in the ecosystem, in AICED (Artificial Intelligence AI), Internet of Things (IoT), Cloud Computing, Big Data, Edge Computing) and other core capabilities essential for digital transformation, and work with operators to expand digital transformation.

5.2 LAN supports diversified operating modes

Demand for one-stop service has been existing in the ICT sector. In the past, communication networks have been constantly trying to provide more services than connectivity. But the reality is, traditional networks still focuses primarily on network layer resources, partly due to technical limitations. On the one hand, wireless cellular networks are scarce resources under continuous wide-area coverage and using good resources to provide good connectivity itself is a very complicated process. On the other hand, wide-area mobile services are featured by complexity, diversity, and difficulty in forecasting, which causes the communication quality fluctuation. Therefore, the information that the network layer itself can provide accurately is limited. In the local enterprise private network scenarios, cellular networks can make significant improvements in network quality management and control, while further optimizing in the resources, and exploring new operating models such as Network as a service (Naas). The network can also proactively detect changes in the application layer, perform data analysis and forecast, and provide relevant services. For example, in a local area private network scenario, not only the layer 1 operation mode(including service KQI, SLA, fault management, user management, policy management, location services, etc.) is provided, but also the intelligent computing-centric layer 2 operation mode（combining the terminal and application layer protocol of local server, with the help of intelligent computing capabilities and content services in local ICT facilities）, to enable the joint optimization of enterprise applications and network layers for the enterprises to benefit from better end-to-end service experience.

6 Preliminary Discussions on Business Model

6.1 Explore new models for serving vertical industries

The premise for profitability of 5G is to provide an end-to-end connection that “changes the society”. In the commercial application of 5G technology, the typical application scenario is to combine with vertical industry applications to expand 2B business model innovations. For operators, the traditional traffic management model can no longer fulfil the value of connectivity and innovative applications in vertical industries. For vertical industries, the evolution of 5G networks to a private network model based on slicing is the general trend. Whether the 5G slicing business model can be established is key to its success in serving vertical industries.

In order to meet the requirements of flexible and diverse business models and give full play to their commercial value, operators can provide different network services through network slicing malls to meet the requirements of various users. The network slicing mall is characterized with customization requirements, enhanced billing, KPI real-time monitoring and open capabilities. Users can choose different SLA slicing/services, slicing capability APIs and user packages in the online slicing mall based on their needs, and order flexibly through the visual interface in one stop. In this way, industry users can participate in network slice orchestration and lifecycle management based on different business models, take advantage of the open capabilities of network slices to build their own end-to-end service processes, and integrate 5G network slices into product design, production, logistics, and sales. In the process, the industry's production efficiency will be greatly improved, the speed of product or service release will be accelerated, thereby reducing OPEX and CAPEX, and improving the E2E user experience. Operators can also resort to this approach to overcome difficulties in 5G technology commercial launch, and truly enable the value of 5G "end-to-end" connectivity, thereby benefiting mobile operators, vertical industry users, end consumer users, etc.

6.2 Promote industry cooperation and create a successful pilot model

The slice-based industry-specific 5G private network can support innovation of multiple business models such as B2B, B2B2C, and B2B2B with 5G as the key technology to respond to different industry application scenarios, bringing opportunities for business expansion of operators and digital transformation of vertical industries. However, presently in terms of business model, product, operation etc., there are differences in perception. It is necessary for vertical industries and operators to work together to build a richer and more diverse ecosystem, so as to build a pilot scenario of 5G intelligent network application in the vertical industries. As more and more achievement models are built, industry users can meet the specific business scenario needs of the industry through the end-to-end slicing private network with guaranteed quality, which helps to boost confidence of vertical industry users in applying 5G technology and enhance their experience. Operators can also accumulate more industry application experience, decompose industry needs in sample points, customize network slices of different service levels, build up network slice template libraries of different service levels, and combine AI and big data technology for KPI real-time monitoring and automatic guarantee capabilities. In general, through the practice of creating pilot models, needs of industry users to customize different network slicing services in different business scenarios can be met, accelerating the process of operators opening network slicing capabilities, strengthening the comprehensive integration of 5G networks and vertical industries to promote the commercialization of 5G technology. Further, the advantages of 5G technology can be introduced to vertical industries which, combined with the characteristics of the specific industries, will benefit the industries in the real sense.

7 Conclusion

As 5G technologies mature, an increasing number of vertical industries are becoming aware of the fact that industry-specific private 5G networks will aid their digitalization and industry upgrade, helping the enterprises meet the needs of a faster and more diverse market. Applications such as telemedicine, AR aid, and machine vision, with the help of 5G network, will profoundly change the ways we live and work.

We also need to be aware of the following three facts in advancing the 5G private networks:

1. Each industry has its own characteristics, making it almost impossible to build a network system that meets the needs of all. Providing slice-based industry-specific private network is the best solution for operators to serve vertical industries, with the benefits of low cost, customization, openness, and professional operation. Enterprises are able to enjoy high bandwidth, low delay, greater connectivity and customization through 5G network slicing, which will propel industrial transformation and upgrading.

2. The B network market is distinctively different from the C network market. Customers of the B network have specific network requirements of "dedicated network", "customized functions", "guaranteed quality", and "agile network construction". Operators need to adjust their organizational structure and working methods for the new business model to ensure that they gain insight in time, track, respond to, and implement customization requirements of customers in vertical industries.

3. The industry needs to increase investment in the terminal field, and promote terminal support for network slicing capabilities and URSP rules as soon as possible, which has become a key part in the 5G service for vertical industry.

In conclusion, the 5G private network service for vertical industries requires dedicated cooperation among operators, equipment vendors, industry application integrators, enterprise customers and terminal manufacturers who need to work together to promote the development and prosperity of the 5G industry.

Acronyms

AMF: Access and Mobility Function 接入和移动管理功能

APN: Access Point Name 接入点名称

BMCA : Best Master Clock Algorithm 最佳主时钟算法

BP : Branching Point 分支点

CAG : Closed Access Group 封闭用户群组

CAPEX : Capital Expenditure 资本性支出

CDN : Content Delivery Network 内容分发网络

CNC : Centralized Network Configuration 集中网络控制

CU : Centralized Unit 集中单元

CUC : Centralized User Configuration 集中用户控制

DNN : Data Network Name 数据网络名称

DS-TT : Device-side TSN translator 设备侧 TSN 转换器

EPC : Evolved Packet Core 演进分组核心网

GM : Grand Master 主时钟

gPTP : Generalized Precision Time Protocol 广义精确时间协议

KQI : Key Quality Indicators 关键质量指标

LADN : Local Area Data Network 本地数据网

MDBV : Maximum Data Burst Volume 最大数据突发量

NaaS : Network as a Service 网络即服务

NEF : Network Exposure Function 网络开放功能

NF : Network Function 网络功能

NFV : Network Function Virtualization 网络功能虚拟化

NPN : Non Public Network 非公共网络

NW-TT : Network-side TSN translator 网络侧 TSN 转换器

OPEX : Operating Expense 营运资本

PCF : Policy Control Function 策略控制功能

PDB : Packet Delay Budget 分组延迟预算

PDU : Protocol Data Unit 协议数据单元

PLMN : Public Land Mobile Network 公共陆地移动网

SBA : Service Based Architecture 服务化架构

SDN : Software Defined Network 软件定义网络

SLA : Service-Level Agreement 服务等级协议

SMF : Session Management Function 会话管理功能

SNPN : Stand-alone Non-Public Network 独立部署的非公共网络

SPN: Slicing Packet Network 切片分组网

TDM : Time-division Multiplexing 时分复用

TSC : Time Sensitive Communication 时间敏感通信

TSN : Time Sensitive Network 时间敏感网络

UDM : Unified Data Manager 统一数据管理平台

ULCL : Uplink classifier 上行分流器

UPF : User Plane Function 用户面功能

URSP : UE Route Selection Policy

Edited by

China Mobile

Huawei Technologies Co., Ltd.

ZTE Corporation

Ericsson Communications Limited

Shanghai Nokia Bell Co., Ltd.

China Datang Group Co., Ltd.

Beijing Baicaibang Technology Co., Ltd.

Lenovo Group Co., Ltd.

Reference

[1] 5G Non-Public Networks for Industrial Scenarios white paper, 2019/7.

[2] China Mobile White Paper on Edge Computing Technology System 2018/6.

[3] 3GPP TS 23.501: "System Architecture for the 5G System".V16.2.0.2019

[4] Analysis of Key Technological Evolution of 5G Core Network for Vertical Industries, Wang Dan, Sun Tao, Duan Xiaodong, Li Yongjing, Liu Tangqing

[5] 3GPP TS 22.104:Service requirements for cyber-physical control applications in vertical domains.V17.1.0.2019.

[6] 3GPP TS 23.502: "Procedures for the 5G System (5GS)". V16.2.0.2019.

[7] GTI 5G MEC White Paper，V1.0.2019