5G Network Slicing for Vertical Industries

Prepared by GSA with contributions from Ericsson, Huawei and Nokia

September 2017

Table of Contents

1. Digital Transformation

As the world becomes ever more digitally and globally connected, industries are experiencing an ICT-driven transformation.

There are many different global trends that can account for this dramatic projected growth, including:

- A rise in emerging economies, which are yet to be fully immersed in the digital world
- The urbanization of integrated mobility
- The digital revolution
- New health and wellbeing demands
- The scarcity and stability of resources.

Additionally, business trends such as hyper competition, new customer power and sophistication, the fast-paced change in business ecosystems and disruptive technological advances all affect vertical industries to different extents.

Figure 1: Key Industries and their chalenges

MANUFACTURING

- Hypercompetition with no sustainable competitive advantages
- Increasing volatility from business cycles and product lifecycles
- The smart factory is advancing from developments in the Internet of Things and automation

HEATHCARE

- Increasing consumer attention on wellbeing
- Increasing cost to fit with social demographic changes
- Increasing demand on quality, patient safety and data storage
- Changing consumer behaviour, freedom of choice and alternative service providers

MEDIA AND ENTERTAINMENT

- Shifting consumer role as a co-creator of media content
- Increasingly interactive and immersive forms of entertainment
- Expansion of digital content through new platforms and new market players (OTT and VOD)
- Ecosystem complexity

FINANCIAL SERVICES

- Disruption from Fintech (technology used to support financial services) due to online payments, e-wallets etc.
- Changing customer relations with online/mobile transactions and customized financial solutions
- Structural changes - state involvement, protectionism and fiscal measures

PUBLIC SAFETY

- Growing public surveillance with CCTV and wearable cameras
- Cyber-attacks - global integration and the digital economy
- Engaged and connected citizens - Internet of Public Safety

AUTOMOTIVE

- Autonomous driving and connected traveler with telematics
- Car sharing and changing commuter habits
- Electric mobility with decreasing battery costs and a green agenda
- Digital enterprise and connected supply chain
- Digial vehicle ecosystem

PUBLIC TRANSPORT

- Infotainment on the move
- Urbanization and Intermodality
- Environmental awareness - $C0_2$ emissions and public spaces
- Urban lifestyle and growing expectations on public transport

ENERGY UTILITIES

- Oil supply imbalance and instability, fracking advancements and carbon constraints
- Structural shifts with increasing retiring assets
- New decentralised business models
- Electrification and renewable energy generation

5G will be a major technology in growing industrial digitalization, creating and enhancing industry digitalization use cases such as immersive gaming, autonomous driving, remote robotic surgery and augmented reality support in maintenance and repair situations.

Figure 1 outlines eight key industries; each of these identified areas faces numerous challenges due to industry trends, which could be addressed with the adoption of 5G digitalization.

Vertical industries have addressed their connectivity and communication needs with dedicated or industry specific solutions.

5G technology will provide a common base to provide a more cost efficient, open, interoperable and large eco-system enabled solution platform for the various vertical industries.

5G is addressing the more stringent and business critical requirements of the vertical industries, such as real-time capabilities, latency, reliability, security and guaranteed Service Level Agreement (SLA)'s.

5G will provide an industry vertical optimized platform catering in an economical way the various requirements and business needs of each vertical. For service providers to offer these capabilities to vertical industries in an attractive, scalable and economical way, they will utilize cloud platforms, analytics, system automation and network slicing technologies as well as new business models.

Network slicing is one of the key capabilities that will enable flexibility, as it allows multiple logical networks to be created on top of a common shared physical infrastructure. The greater elasticity brought about by network slicing will help to address the cost, efficiency, and flexibility requirements imposed by the large variety of industrial vertical services. Moreover, network slicing will help new services and new requirements to be quickly addressed, according to the needs of the industries, i.e. a faster Time to Market.

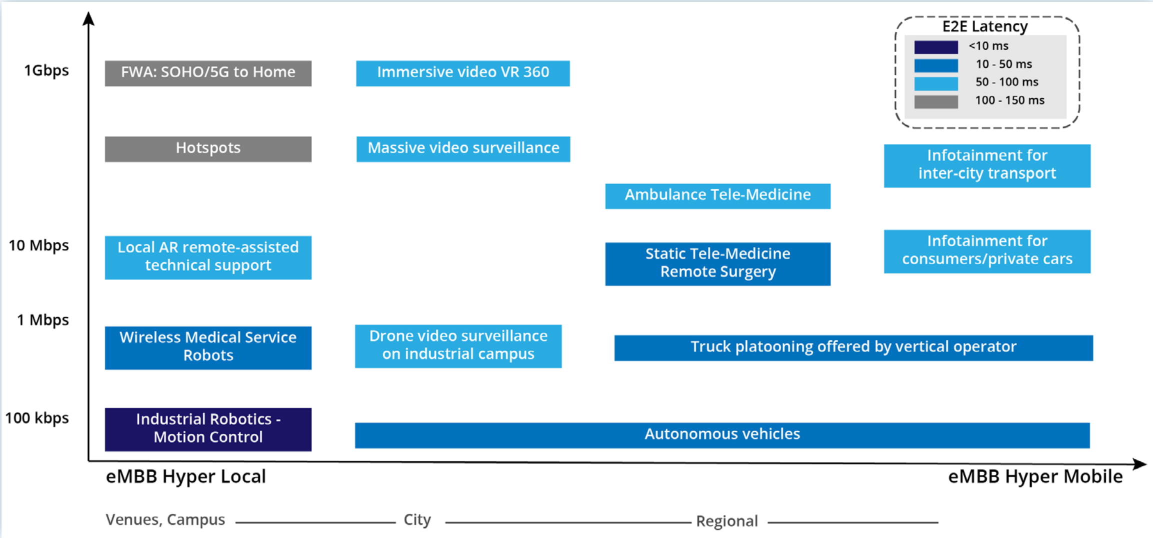
2. Vertical Requirements

5G will allow the creation of logical networks on top of a shared infrastructure. Those networks can be optimized for specific use cases, service types and support various business models. "Network as a service" is the tool for implementing dedicated and customized virtual end2end networks, enabling vertical industries to rapidly deploy their services.

Vertical industries are very diverse and their requirements are dictated by the service characteristics of the related vertical segment. Examples of services requiring Low Latency and High Reliability are Autonomous Vehicles, Industrial Control, and Augmented Reality. Figure 2 on the next page shows the most relevant segments.

Requirements can be clustered in 'Operational Requirements', 'Functional Requirements' and 'Performance Requirements' categories. These three categories can be interpreted via the Maslow model, which indicates that some vertical industries have a higher hierarchy of requirements than the others. In order to catch such diverse requirements, flexibility is one of the key feature for 5G system, Figure 3.

Figure 2: Service Requirements by vertical segments



2.1 Category 1: Performance requirements

> *"Global operators, vendors, and vertical industry partners should work together to facilitate the development of network slicing in terms of application scenarios, technologies, standards, and the industry value chain." -- Huawei*

Latency:

The target of very low latency combined with high reliability of 5G will enable many applications to be untethered, supporting a larger number of use cases that today require fixed line connectivity. This is the case for industrial applications, where requirements for end-to-end latency cover a range between 1000 ms to below 5ms, the latter for mission-critical applications where real-time transmissions are required for motion control.

Throughput, Up Link (UL) and Down Link (DL)

In the future Mobile Broadband services will go far beyond basic mobile Internet access and will cover rich interactive work, media and entertainment applications in the cloud or reality augmentations (both centralized and distributed). While this is already happening to some extent, the increase in video resolution, widespread use of 1080p, introduction of 4K to wireless and future with 8K, generate the need to support increasing high levels of throughput especially in the DL. Additionally, this is also affecting the UL, which supports

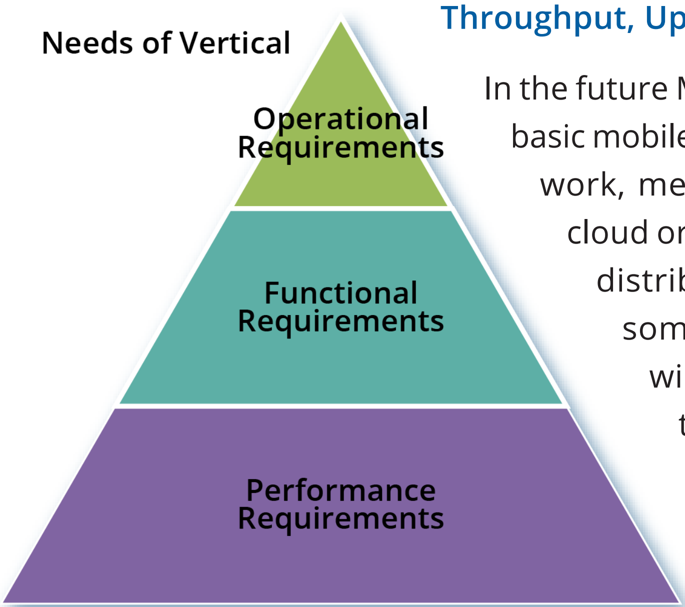


Figure 3: Vertical Industry Requirement Maslow Model

better performances compared to LTE, for use cases such as video surveillance and remote control. A key consideration is that intelligent machines produce huge data volumes per second as a result of their multiple embedded sensors. These sensors can generate everything from small data packets up to very high broadband data. The aggregated raw data can be some hundreds of megabytes per second which cannot be shared in an easy way with other machines or operation centers without being pre-processed and filtered before sending over the wireless modem.

Depending on the service and scenario, 5G can offer user data rates in the order of Gbps in specific scenarios, like indoor and dense outdoor environments. Data rates in the order of tens of Mbps should be accessible everywhere, included sparse populated areas.

Availability and Resilience (service always availability)

High availability is essential to ensure minimal service accessibility to critical infrastructures or service providers in case of a disaster. This is achieved by extending the network coverage, ensuring network access and providing roaming support to make sure that critical applications are not in outage when needed. For some emergency communication services, e.g. public safety, hospitals, 99.999% availability is required (i.e., less than ~300 seconds of accumulated outage per annum), aiming at equating availability of wireless and wired solutions.

Resilience deals with the capability of the network to recover from failures, and it is an essential component for maintaining high availability as well as high data rates.

Reliability

Reliability is defined as the probability that a certain amount of data to/from an end user is successfully transmitted to another peer within a predefined time frame, i.e., before a certain deadline expires. The amount of data to be transmitted and the deadline depend on the service characteristics of the underlying use case. For some category of services, like "Factory Industrial Process automation & Motion Control" and Transport systems, a reliability of at least 99.9999% is requested; this means that the probability of a packet not delivered within the specified deadline is to be below 10^{-6} (or equivalently, at most one in a million packets does not arrive within the specified deadline, on average).

In 5G several multi connectivity techniques, such as Coordinated Multi-Point (CoMP), enhanced Carrier Aggregation and Dual connectivity (data duplication) as well as simultaneous Radio Access Technologies (RATs) (e.g. 3GPP, non-3GPP, fixed) could be exploited to achieve the desired reliability.

Coverage

While some industries need services only in a local area, other many require a full coverage. For example logistics and freight tracking need wide coverage and reliable location information for inventory and package tracking wherever they are.

Other services have machines geographically located in many confined area, for example sensors/actuators in a hospital, refinery or garage. The 5G network will be able to efficiently manage several hot-spots areas which might be located in challenging coverage-positions (e.g., indoors or at the cell-edge). Furthermore, the service provider may not want to be dependent on the coverage and rollout strategy of a given operator and may therefor want to provide coverage on its own.

2.2 Category 2: Functional requirements

Security and Identity Management:

Secure communications need to guarantee that personal or confidential data must not reach the public domain and not be modified or replayed by unauthorized parties.

Verticals industries will require different level of security. For example security mechanisms used for ultra low-latency, mission-critical applications (e.g. autonomous driving, control of a smart grid or smart operation of industrial automation processes) require a high level of communication security and may not be suitable in massive Internet of Things (IoT) deployments where mobile devices are inexpensive sensors that have a very limited energy budget and transmit data occasionally.

Identity Management: relates to the management of devices and user identities. 5G will connect a huge variety of devices and users which have subscriptions to vertical service providers. This demands for a new device-user identity management and related lifecycle management, which will complement the universal SIM (USIM). This new mechanism is required because devices such as sensors, smart home and wearable devices are too small or cheap to host a USIM card, which is the way used today in cellular networks to manage user identities as well as keys.

User-device security: different mechanisms for identifying and authenticating devices and/or their subscriptions needs to be supported:

- Authentication by the network only: service providers may rely on network authentication only saving costs for managing service authentication
- Authentication by service providers only: authentication is performed from service provider only exempting devices from radio network authentication
- Authentication by both network and service providers, each for the related domain. This is the model adopted today
- Protect control and user plane by means of encryption at network layer and or application layer. In addition applications may require not only encryption, but user plane integrity protection service thay is able to guarantee the privacy of all users' communications.

Isolation

While using a shared network infrastructure, the different vertical industries (and optionally the various services), need to be isolated one from another. That means that each virtual network belonging to different vertical customers are protected, preventing their resources from being accessed by network nodes of others. This is necessary to ensure a reliable and warranted service assurance, together with data and communication integrity and confidentiality.

2.3 Category 3: Operational requirements

Verticals will be able to timely deliver their service to their own customers by ordering a network slice it in a simple way, e.g. through a simple user interface. Several industries may rely on the Mobile Network Operator (MNO) for the deployment and management of their own network slice while other industries may want to do it independently.

Self-Management of resources/policies, APIs, Service Assurance for core requirement of vertical's business

Depending on assets and business, multiple models are needed for managing service/applications and customers of the verticals.

In one model the vertical industry customer has their own customers but no network resources. The vertical requests a network slice via a North Bound Interface (NBI) of the MNO, realized e.g. though a dashboard-like web service.

The network slice is chosen and ordered from existing templates by filling the service description and related parameters. KPI targets are as well monitored to assure SLA conditions for the delivered service.

Another model is the one of a vertical who owns part of network resources and designs and customizes own service. In this case the vertical may need part of the network (for example the RAN) of the MNO according to a given SLA via the NBI as in the case above.

Charging/ Billing

Verticals might have different requirements for charging - ranging from aggregated network usage information collection provided by a telco operator up to detailed per subscription accounting information. The information collected should be suited to check against SLAs the vertical has with the MNO in terms of services, quality of services, overall performance and the resources assigned.

Different requirements apply also to the interface for charging data delivery ranging from usage collection/recording to the supervision of the execution in real time.

Global Operation

Vertical industry customers are seeking solutions for digital transformation, meanwhile, they also expect that digital transformation could also bring new opportunities, e.g. business globalization. Enterprises such as manufacturing normally have factory plants as well as business branches spread around the world. To achieve overall connectivity of all the business units they will optimize the operation process as well as logistic costs. Global operation therfore requires connectivity that crosses different administrative domains.

"Slicing provides the tools to effectively address new business opportunities where connectivity will enable transformation of industries and society. Where previously such opportunities were addressed by building multiple dedicated infrastructure solutions, with slicing, these network solutions can be provided by software on a common network infrastructure. Network virtualization and slicing represent a similar shift as when dedicated computing hardware were replaced by virtualization and cloud computing", -- Håkan Andersson, Director Industry Alliances, Ericsson

3. 5G Network Slicing

3.1. Definition

Traditional telecom networks were built with physical nodes only, comprising one "monolithic" network that was to provide all the services offered by the network operator. The network was statically configured, mostly manually managed and a "single tenant environment".

Driven by technology evolution as well as business needs, the way that telecom networks are built is changing in order to cope with new demands. Key technologies are virtualization, Software Defined Networks (SDN), orchestration capabilities, advances in air interface along with a general evolution of software technologies.

With these technologies at hand, it becomes possible to view the telecom network in a new way. Logical networks can be designed, instantiated and operated in a dynamic "on-demand" basis, targeting the specific needs of specific customers, services or business segments, be it related to service characteristics, business models, self-management of resources/service etc.

Such "logical networks" are denoted "network slices". To be more concrete, a network slice comprises of a group of network functions, resources and connection relationships. It typically covers multiple technical domains, which includes terminal, access network, transport network, core network, data center domain that hosts third-party applications, as well as network management system, etc..

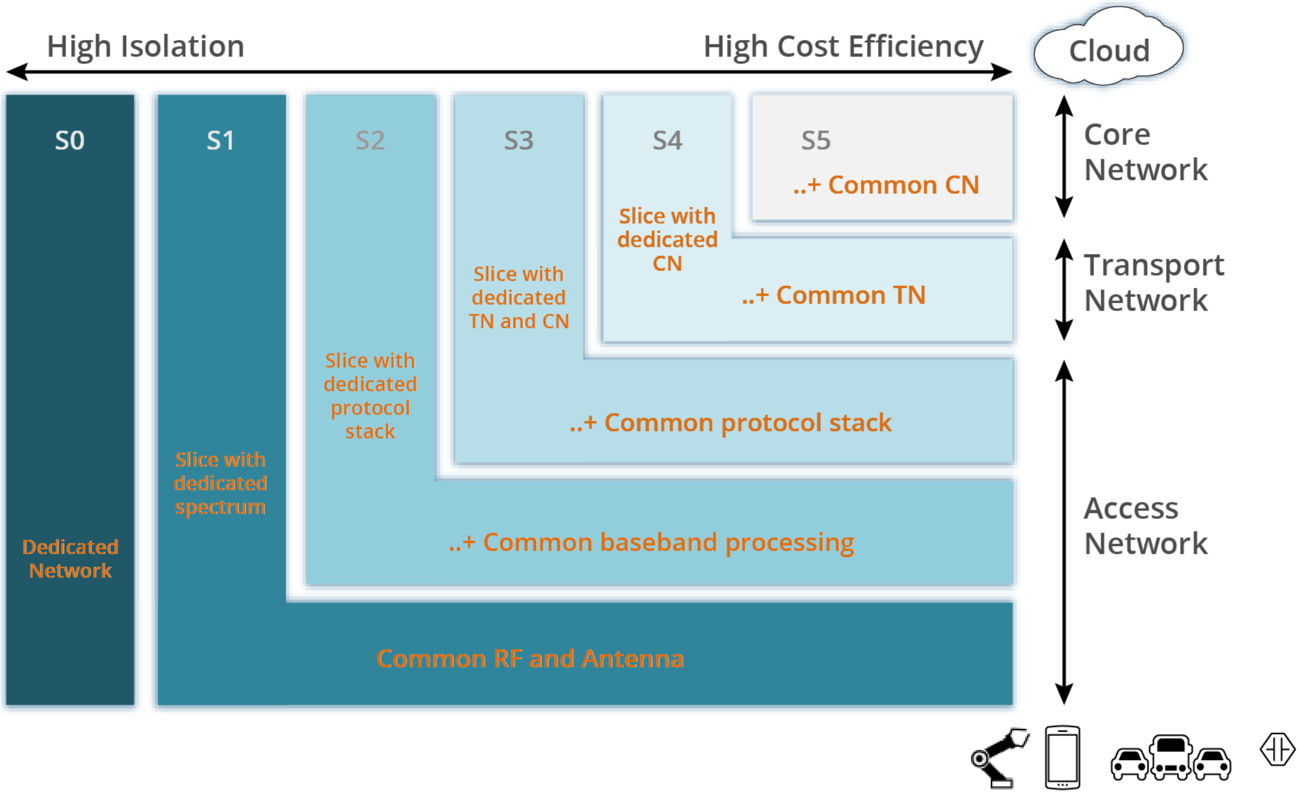
3.2. Network Slicing Forms

As the different business contexts or customers which they serve are very diverse, network slices will need to come in very different shapes and forms. This calls for a very high degree of flexibility and agility. Therefore, also as part of the network slicing concept comes a much a higher degree of automation, orchestration and advanced service creation capabilities.

For instance, a network slicing concept could penetrate all technical domains of the communication system in order to tackle different categories of requirements from verticals. For the access network, from lower level spectrum to higher level protocol stack, it could be sliced and tailored especially to handle performance requirements, e.g. a protocol stack could be designed and optimized for Ultra-Reliable Low latency Communications (URLLC) purpose.

For the transport network, isolation could be achieved either via resource sharing (e.g. virtualization) or dedicated resource (e.g. dedicated port and cable for a network slice provided for public safety service). The core network could provide customized network functions upon the requests from the vertical industries, e.g. simplified mobility management function for massive Machine Type Communications (mMTC), or a complicated mobility management function to handle ultra-high speed terminals on high way for the automotive industry, as can be seen in Figure 4.

Figure 4: Network slicing different solution spaces in order to handle diverse vertical requirements



All of this represents a paradigm shift within the telecom industry that is believed to be essential for the 5G ecosystem. This paradigm and transformation journey can be referred to as “network slicing”.

3.3. Tackle Vertical Requirements via Network Slicing

Network slicing will play a significant role in addressing the vertical requirements described in section 2. The ways in which that can be achieved is described below:

Category 1: Performance requirements:

Latency:

Different network slices can use different distribution models of network function and applications, optimizing the topology through which the traffic traverses to match the needs of the customer. Furthermore, it is possible to couple this to resource reservations, separate paths etc. at the infrastructure level that brings additional improvements. In the "monolithic" model of the past, latency was the result of a tradeoff (and competition for resources) between all the service types.

> *"Adaptable bandwidth on demand in highly dense areas like stadiums, can scale accordingly to user demand.*
>
> *Having flexibility and control of operational resources is essential to cope with variable and changing demands in business operations" -- Nokia*

Throughput, UL and DL

General throughput increase comes from technology evolutions in the access and infrastructure layers. Network slicing however makes it easier to "allocate" the capacity according to need. A network slice where high throughput is needed can for example be designed to have "fat" user plane pipes, while other slicers do not. Again, in the past model, tradeoffs had to be made.

Availability and Resilience (service always availability)

High availability and resilience typically comes at the price of more resources needed, which translates into higher cost. In a "monolithic" network, all services must use the same approach to availability and resilience, which means that either some services get penalized in terms of higher cost than required, or some services get lower availability and resilience than required.

With network slicing, different network slices can have levels or even types of availability and resilience. Instead of poor compromises it instead becomes possible to customize and optimize each service.

Reliability

With network slicing, network functions and resources can, to a much higher degree, be isolated and reserved. Although 100% reliability can never be guaranteed in real life, network slicing brings a higher level of control and predictability.

Coverage

Coverage ultimately comes from the access infrastructure (which needs to be there). Network slicing will allow a more flexible control of how the resources are assigned to and used by different services.

Category 2: Functional requirements:

Security and Identity Management:

Network slices are thought of as "isolated". Although this is a relative term, this also means that the level of isolation of a network slice can be customized, just like many other things. For example:

- A network slice could, if desired, be cut off from the Internet and only connecting to private VPNs
- A network slice can be equipped with more advanced security protection, e.g. enterprise grade firewalls
- A network slice could, if needed, be given separate physical resources
- As a network slice has its own logical network functions, there will be fewer "attack vectors"
- A network slice can use data that is isolated in relation to other network slices.

Identity Management/User device security

Due to the fact that network slices are isolated, it becomes possible to use other forms of credentials or user data basis within a network slice. In the past, this could have posed a threat on the common infrastructure and on all services (as they were shared).

Category 3: Operational requirements:

Self-Management of resources/policies, APIs, Service Assurance for core requirement of vertical's business

As network slices can be designed to have a large degree of isolation (in relation to other network slices), it becomes significantly easier to allow customers to self-manage resources and policies (as it can be done in a "safe" way, not impacting any other customers or their resources.

Furthermore, network slices can be highly customized in terms of functionality, allowing for instance different forms of service assurance capabilities to be active within one slice as compared to others. This way, the specific requirements can more easily be addressed.

As slices are independent and isolated, the functional content and resources assigned can be very different, which again allows for better customization and matching the needs of the customer.

Charging/ Billing

As network slices are isolated and use their own instances of network functions, charging systems or even billing systems, it is also here much easier to customize towards the specific needs of a customer or a business model. In the past model, such customization would typically require a large degree of time consuming system integration.

Global Operation

The network slicing concept can be extended globally in different ways. One way is that a customer gets service from multiple service providers, each having their own slice (all realizing the same service). A simpler model for the enterprise customer may be that one service provider (operator or some intermediary) delivers the service by realizing one network slice with global reach. This service provider integrates services from a set of partners (e.g. in different countries) acting as sub-suppliers, contributing to building the global reach of the service.

4. Business Use Cases

As discussed in Section 1, digital transformation will influence a wide range of industries. In section, two representative use cases are described below. Such use cases could be developed for all relevant verticals.

4.1. Automotive Industry

Vehicle communication has been investigated long before 5G. It is possible to perform automatic driving via in-car sensors like camera, radar, lidar. However, 5G's design is targets to offer stringent performance, such as ultra-low latency and ultra-high reliability that will bring vehicle communication to the next level, e.g. from assisted-driving to cooperative autonomous driving.

> ***"Adaptive Network slicing and predictive Network performance information will drive highly automatic and autonomous driving to the next level of comfort and performance" -- Volkswagen***

Vehicle communication that relies on a network requires a wide coverage range, e.g. along the highway, in rural or in the urban area. Hence, the operator will play a vital role in the ecosystem of the automotive industry. Due to security concern, automotive companies may deploy their private cloud environment to host automotive applications. In order to cope with performance requirements, it is essential for a vertical customer to seamlessly integrate its resources as well as services together with an operator's territory.

4.2. Industry 4.0

Due to the advanced progress of IoT, industrial automation, cloud technology, etc., manufacturing industry is also undergoing the digital transformation process. Being different from above mentioned automotive industry, players from manufacturing industry normally have more diverse business roles in its own ecosystem. On one side, they could be providers for external customers in terms of mechanical components, technical solutions, etc., and on the other side, they could be the users of their own solution, e.g. digitizing their own plants. Especially for the cases like control of industrial machines, communication is normally limited within the factory campus. Manufacturing industry players may well be using private networks (unlicensed or granted usage of licensed spectrum) tougher with their own cloud resources.

The business scope for such manufacturing industry is very international. For tier-one players, the amount of plants may be in the order of two to three digits and spread in different countries and even continents. Therefore, global network slicing operation is essential for such use case. Moreover, global network slicing concept will also bring added-value for the vertical customer in terms of business development, because under one unified platform could shield the complexity of business negotiation among different administrative domains.

5. Summary

From 1G to 4G, mobile communication is constantly changing our behavior, experience, and life style. Coming to the 5G era, the mobile communication will further change our society by providing solid foundations to realize the "Internet of Everything": connected sensors to enable a smart city, connected robots for manufacturing plants, human mobility improved by connected cars and public transportation, etc. To enable such a vision it is essential to have a deep understanding of the various requirements of the vertical industries and serve them without sacrificing efficiency while keeping costs low. As one of the key features for 5G, network slicing allows a flexible system architecture able to deal with those requirements. 5G is a communication vision turning point, which will have deep impacts on telecommunication as well as vertical industries. Hence, the engagement from all parties are essential to make 5G a successful story, and make industry digitalization a reality.

Glossary

ACRONYM	DEFINITION
API	Application Program Interface
CoMP	Coordinated Multi-Point
DL	Down Link
IoT	Internet of Things
KPI	Key Performance Indicator
mMTC	Massive Machine-Type Communications
MNO	Mobile Network Operator
NaaS	Networks as a Service
NBI	North Bound Interface
RAT	Radio Access Technology
SDN	Software Defined Networks
SLA	Service Level Agreement
UL	Up Link
URLLC	Ultra-Reliable Low latency Communications
(U)SIM	(Universal) Subscriber Identity Module
VPN	Virtual Private Network

About GSA

GSA (the Global mobile Suppliers Association) is a not-for-profit industry organisation representing companies across the worldwide mobile ecosystem engaged in the supply of infrastructure, semiconductors, test equipment, devices, applications and mobile support services.

GSA actively promotes the 3GPP technology road-map – 3G; 4G; 5G – and is a single source of information resource for industry reports and market intelligence. GSA Members drive the GSA agenda and define the communications and development strategy for the Association.

Membership of GSA is open to any supplier of products; systems or services related to the mobile industry and brings many benefits including access to the GAMBoD database. The range of benefits includes enhanced discussion, networking and influencing opportunities on the key industry topics, and unique promotional/visibility opportunities for your company name, capabilities, positioning and messages. More details can be found at https://gsacom.com/gsa-membership/

Website http://www.gsacom.com

News/updates RSS Feed: https://gsacom.com/rss-feeds/

GSA LinkedIN group: www.linkedin.com/groups?gid=2313721

Twitter: www.twitter.com/gsacom

Facebook: www.facebook.com/pages/Global-mobile-Suppliers-Association-GSA/123462771012551

NOTES:

Errors & Omissions Excepted

Contact

Contact: GSA Secretariat

Email: info@gsacom.com
Tel: +44 330 113 1572

GSA website: www.gsacom.com