Uen 284 23-3244 | January 2015

ERICSSON

5G SYSTEMS

ENABLING INDUSTRY AND SOCIETY TRANSFORMATION

A digital transformation, brought about through the power of connectivity, is taking place in almost every industry. Through an unprecedented ability to share information, people and industries are collaborating more, creating solutions that combine many different areas of expertise and overturning traditional business models.

But the pressure on the networks that provide this connectivity is palpable. For that reason, 5G systems will be built to enable logical network slices, which will enable operators to provide networks on an as-a-service basis and meet the wide range of use cases that the 2020 timeframe will demand.

INTRODUCTION

Similar to the way in which the music industry has overhauled its business model, other industries are changing the way they do things, moving away from physical consumption toward virtual products, adopting smarter ways of working and collaborating at a greater level. This transformation places new requirements on connectivity and sets the scene for the next generation of wireless access – 5G – systems that are set for commercial availability sometime around 2020.

The use cases predicted for the 2020 timeframe will need new types of connectivity services that are highly scalable and programmable in terms of speed, capacity, security, reliability, availability, latency and impact on battery life.

And so, traditional cellular networks and their one-size-fits-all approach need to be adapted so that the thousands of use cases, many different subscriber types and varying app usage can be supported.

Evolving existing radio-access technologies (RATs), like LTE, and new 5G technologies will all be part of a future flexible and dynamic 5G system. Support will exist for cross-domain integration and multi-RAT environments. New technologies will enable concepts like very low latency, and the need for additional capacity will require communications to operate in higher frequency ranges than they do today.

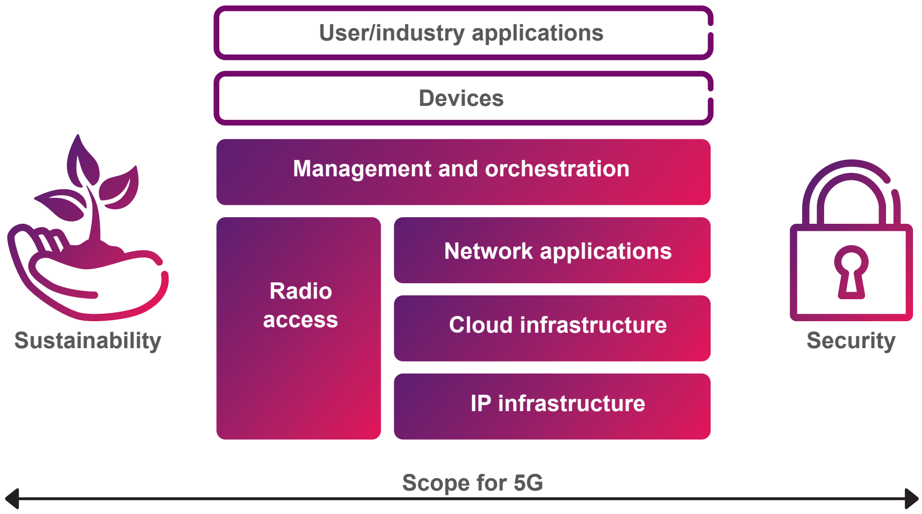

Figure 1: Scope of 5G systems.

Ultimately, the evolution to next generation 5G systems will enable operators to provide networking as a service.

USING 5G SYSTEMS

There is a general industry consensus regarding the increase in demands that 5G systems will need to meet in comparison with today's networks. This common understanding indicates that traffic volumes will be multiplied 1,000 times; 100 times more devices will require connectivity; some applications will demand data rates 100 times the speeds that average networks currently deliver; some will require near-zero latency; and the entire system will work to enable battery lives of up to 10 years.

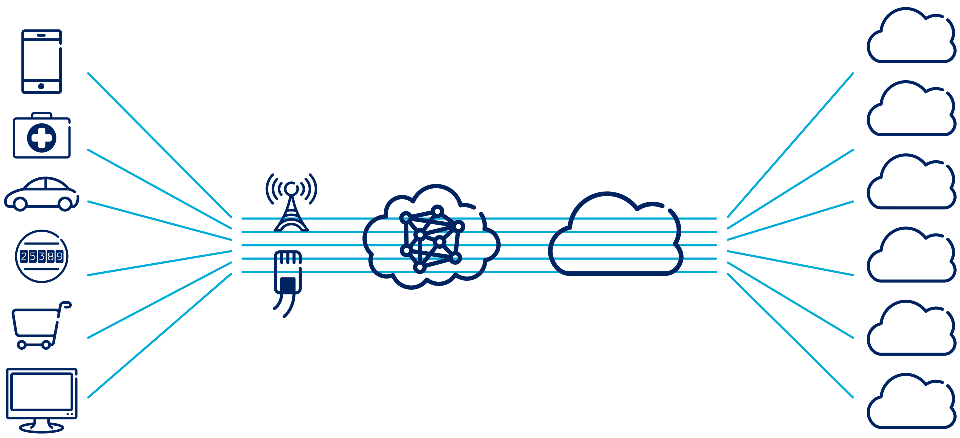

Figure 2: One network – many industries, many network slices.

But not all future use cases will require networks that are ultra-fast, super smart, and have the capacity to support massive numbers of devices. Instead, networks will be built in a flexible way so that speed, capacity and coverage can be allocated in logical slices to meet the specific demands of each use case.

For example, the best connectivity service for remote operation of heavy-duty machinery has low latency, high bandwidth, covers a limited geographical area and supports a known maximum number and type of devices.

Using today's models, a construction or mining company needing to operate machinery remotely might decide that the best option is to create such a connectivity service by building an owned network and then tearing it down when it is no longer needed. However, in the future, 5G systems will enable operators to offer such niched connectivity as a service.

Video is also an important aspect of network evolution. This segment already accounts for a significant part of overall network traffic, and it is expected to grow around 13 times by 2019 – by which time forecasts estimate that it will account for over 50 percent of all global mobile data traffic [1].

As 5G systems will be multi-RAT networks, combining LTE and new 5G technologies, the resulting flexibility in radio resource utilization will enable prioritization of media content. This capability implies that 5G systems are likely to become the preferred means of content delivery for personalized media.

To meet the demands of such widely varied use cases, 5G systems will, therefore, be built with technologies that use logical instead of physical resources, and will enable operators to provide networks on an as-a-service basis. Such network services will provide the flexibility needed to allocate and reallocate resources in line with demand, and to tailor network slices to specific needs.

NETWORK SLICES

Traditional, one-size-fits-all network architectures with purpose-built systems for support and IT worked well for single-service subscriber networks with predictable traffic and growth. However, the resulting vertical architecture has made it difficult to scale telecom networks, adapt to changing subscriber demands and meet the requirements of emerging use cases.

Cloud technologies together with software-defined networking (SDN) and Network Functions Virtualization (NFV) provide the tools that enable architects to build systems with a greater degree of abstraction – which enhances network flexibility. Cloud, SDN and NFV technologies allow vertical systems to be broken apart into building blocks, resulting in a horizontal network architecture that can be chained together – both programmatically and virtually – to suit the services being offered and scaled.

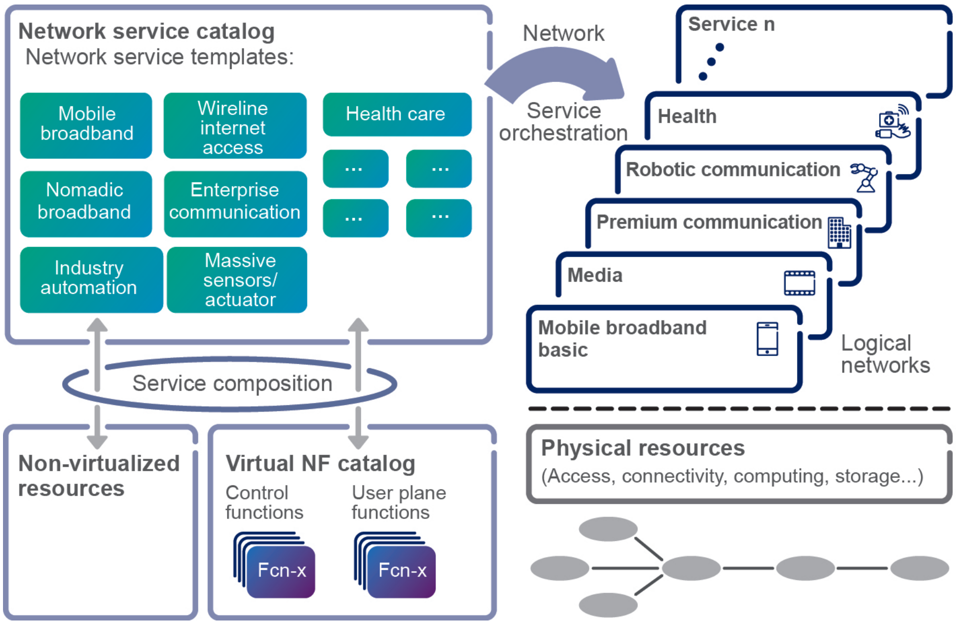

Figure 3: Service creation using network slices.

In 5G systems, networks will be further abstracted into network slices: a connectivity service defined by a number of customizable software-defined functions that govern geographical coverage area, duration, capacity, speed, latency, robustness, security and availability.

The concept of network slices is not a new one; a VPN, for example, is a basic version of a network slice. But the wide range of use cases and tougher requirements that future networks will need to support suggests that network slices in the context of 5G will be defined on a whole new level, more like networks on-demand.

NETWORK SLICES IN THE CONTEXT OF 5G

The benefit of slicing networks is not just the capability to deliver a wide range of connectivity services to any industry, but also to ensure that usage can be billed accordingly. Slicing networks provides greater insight into network resource utilization, as each slice is customized to match the level of delivery complexity required by the service or services using the slice.

SMART METER SERVICE

A utilities company, for example, requires connectivity for its smart meters. This need could be translated into a network slice that connects a number of machine-to-machine (M2M) devices with a latency and data rate that enables download of status updates within a given time. The security level of the service is medium, and it is a data-only service that requires high availability and high reliability.

Additional network functions associated with higher levels of security, longer durations or increased reliability could be configured to suit the needs of the application.

SENSOR UTILITY SERVICE

The same utilities company may require connectivity for its fault sensors. The network slice for this type of service needs to be able to receive round-the-clock status indicators from all the M2M devices in the system. This use case requires data-only coverage with high availability and robustness, and medium security and latency.

Again, depending on the use case, a network slice delivering this connectivity service can be configured with different network functions to enable higher levels of security, or near-zero latency, for example.

MEDIA SERVICE

An operator wants to provide all of its subscribers in a given country with very high throughput for its video-streaming service. High data speeds and low latency are the key requirements for this network slice.

CAPACITY/COVERAGE ON DEMAND

Certain mission-critical services may need instant access to network capacity or coverage in the event of an emergency. Such use cases can be prearranged through agreements and provided by the operator on demand. The critical aspect for this use case is the business agreement.

KEY ENABLERS

The concept of network slices requires a number of technologies to be in place. SDN, NFV and cloud technologies enable networks to be broken out from their underlying physical infrastructures so that they can, programmatically, provide connectivity as a service.

Devices, which will vary greatly in nature, will need to use connectivity in a smart way, with minimal signaling, maximized sleep cycles and lean data transmission – sending only the information needed for the given service.

Security needs to be provided end-to-end (E2E), so that services mashed together from multiple sources can be trusted. Management systems need to be adapted to monetize the wide range of device types that will use connectivity. Networks will need to be able to support several RATs simultaneously, and the radio environment will need to support service flexibility, all while maintaining control over costs and energy consumption.

SOFTWARE-DEFINED NETWORKS

The benefit of SDN lies in its ability to provide an abstraction of the physical network infrastructure. Through network-wide programmability – the capability to change the behavior of the network as a whole – SDN greatly simplifies the management of networks.

The level of network programmability provided by SDN allows several network slices, customized and optimized for different service deployments, to be configured using the same physical and logical network infrastructure. One physical network can therefore support a wide range of services and deliver these services in an optimal way.

NFV

By separating hardware from software, NFV allows a network function to be implemented programmatically instead of by a physical piece of hardware. This capability enables instant scalability, which supports the delivery of services like capacity or coverage on demand.

The most significant benefit brought about by NFV is the flexibility to execute network functions independently of location. By virtualizing a network function, it is no longer bound to a specific location or node. The same network function can be executed in different places for different network slices. Depending on the use case, a network function could either be placed in a centralized data center (DC) or close to a base station. By placing network functions accordingly, the same physical infrastructure can provide connectivity with different latencies.

RADIO ACCESS

To provide the wide range of future wireless use cases with customized connectivity, 5G systems need to be based on a flexible radio-access solution that can adapt to different requirements and deployment types.

In terms of spectrum, 5G systems will need to be able to operate over a very wide frequency range from below 1GHz up to and including millimeter wave (mmW) frequencies.

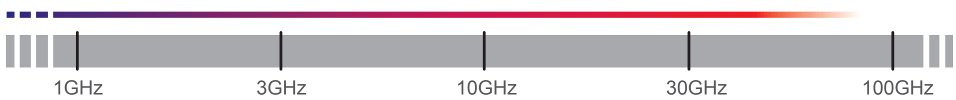

Figure 4: Spectrum range for 5G use cases.

Lower frequencies will remain a key part of the spectrum used to deliver 5G network capabilities such as very low latency, ultra-high reliability and high data rates with wide-area coverage. This range will be complemented by high-frequency deployments – above 10GHz – that will be able to deliver extreme data rates and extreme capacity in dense areas.

Operation under different spectrum-assignment regimes will enable 5G systems to provide greater flexibility. Since its inception, cellular communication has relied on dedicated licensed spectrum, and this will remain the case for 5G systems in order to deliver high-quality connectivity with very high availability. However, the use of dedicated licensed spectrum will be extended to operate with more flexible spectrum assignment. For example, Licensed Shared Access (LSA) will enable spectrum sharing across industries, and licensed-unlicensed spectrum combinations can be operated through Licensed Assisted Access (LAA). By using carrier aggregation technology, LAA enables mobile devices to leverage the combination of both licensed and unlicensed spectrum bands.

In terms of technology, LTE will continue to evolve, and will be an important part of the overall 5G wireless-access solution – providing a backward-compatible path for 5G capabilities in LTE spectrum. In parallel, new RATs will emerge, initially targeting new spectrum where backward compatibility to existing technology is not an issue. New spectrum will primarily be available in the higher frequency bands above 10GHz. However, new spectrum may also partly become available in lower frequency bands. On a longer timescale, new RATs may migrate into LTE spectrum.

A high level of interworking between LTE evolution and new RATs is needed to ensure that 5G functionality can be introduced smoothly. Such interworking will include dual-connectivity where, for example, a device will maintain simultaneous connectivity to a dense high-frequency layer providing very high data rates as well as to an overlaid lower-frequency LTE layer that provides ubiquitous connectivity. User-plane aggregation between LTE and any new radio technology is another example of the high level of interworking.

Additional key technology components for the 5G radio-access solution include:

- advanced multi-antenna technologies
- ultra-lean transmission to maximize resource efficiency
- flexible duplex
- access/backhaul integrations, where access and (wireless) backhaul share the same technology and the same overall spectrum pool
- well-integrated device-to-device communication.

These technology components will not only apply to the new technology part of the 5G wireless-access systems but will, to a large extent, also be applicable to the evolution of LTE.

DEVICES

Early cellular devices were primarily designed to support voice services. Today's devices support instant messaging and social networking, as well as general access to and sharing of content, information and media. Device types have evolved from being voice-centric to high performance internet-access mobile devices including laptops, smartphones, tablets and wearables.

Consumer electronics are increasingly becoming connected, and connected devices are becoming more commonplace in the fabric of things and machines – everything from cereal boxes to cars, homes and agriculture.

Consequently, 5G systems will need to support a much wider range of devices than networks do today. Network slices can be configured to meet the needs of the entire range of applications supported by these device types, whether they require low latency, minimal impact on battery consumption or wide area coverage. Slicing can therefore provide the necessary support for differentiated service offerings.

E2E SECURITY

An important part of the vision for 5G systems relates to trust in shared resources for computing, communication and storage. Such trust depends on necessary security functionality, such as identity management and pervasive integrity protection, as well as the proper implementation of technology and delivery of services: security assurance.

It is likely that a number of key security technologies will play a major role in 5G systems. Expressing individual security goals for network slices serving diverse needs and users requires Security Service Level Agreements (SSLAs): the ability of the network – from a security and resilience perspective – to provide explicit, measurable and enforceable capabilities enabling the network to behave consistently in a prescribed way. Such SSLAs specify the relevant security constraints for traffic, such as traffic isolation requirements, desired levels of availability or the level of security assurance of network elements.

Using trusted computing technologies rooted in hardware security anchors enables all the equipment comprising the 5G system to be booted in a trusted way. In addition, hardware security anchors provide cryptographically strong E2E security associations, from say a mobile phone, all the way to application services that are executed in cloud DCs. The software version that individual participants are running can be remotely assessed, which allows the correct SSLA configuration to be determined. Security functionality to implement SSLAs should, therefore, be rooted in trusted computing technologies, as they offer the strong level for security guarantees required in 5G systems.

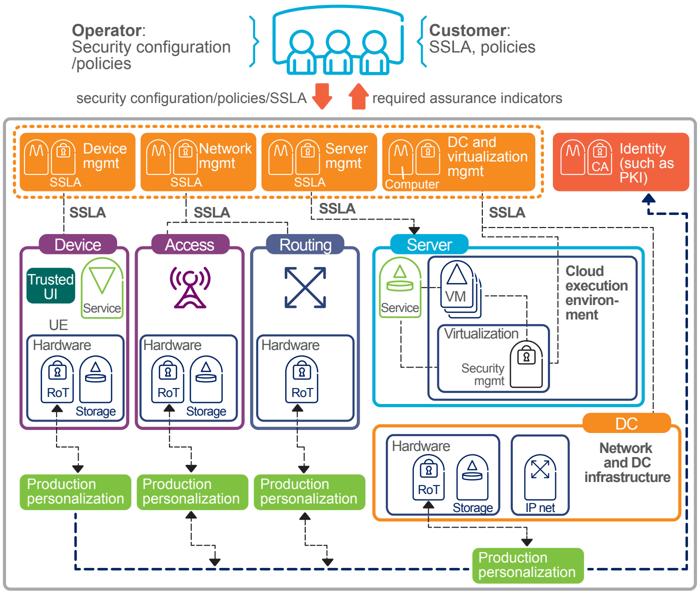

Figure 5: Security in 5G systems.

MANAGEMENT AND ORCHESTRATION

Operators, as well as enterprises, want to deliver and operate services at the lowest possible cost without compromising quality. To do this, management entities in 5G systems will need to be able to automate and orchestrate provisioning processes, as well as be capable of coordinating cloud resources for complex dynamic systems that require resource access control and service quality management. Given the increasing demand for Service Level Agreements (SLAs), the network needs to play a more active role in delivering reliability and performance guarantees.

Orchestration enables automation across the building blocks of a network through centralized management of network resources.

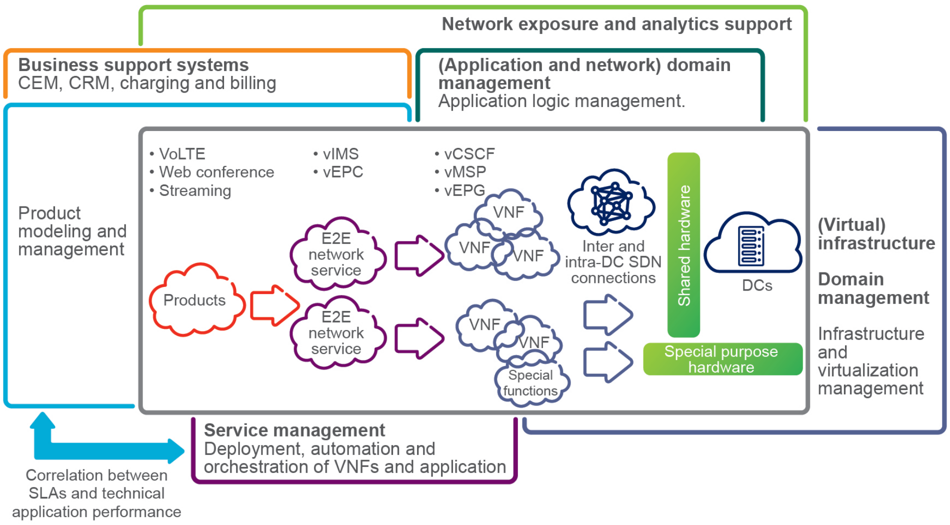

Figure 6: The management landscape.

To offer services that draw resources from several building blocks, E2E orchestration is needed to match external business offerings with network efficiency. To optimize media delivery, for example, orchestration would place virtualized network functions on resources that are physically close to the subscriber.

New technologies will be needed so that 5G systems will be able to cope with the increased load. This includes technologies and methods for improving performance and operation of the 5G system through data analytics, including:

- enhancement of RAN features based on data analytics
- automation – self-organizing networks and data analytics features for operations support systems and network management systems
- automation and data analytics technologies for network operations centers.

ENERGY PERFORMANCE

Networks need to be energy-focused on two fronts: reduced energy consumption of all network resources to lower operational cost, and increased battery life to enable massive deployment of sensor networks.

Wide area coverage dominates network energy consumption, and studies [2] [3] show that for LTE, approximately 50 percent energy savings are possible without compromising coverage and while still meeting the increased – 1,000 times – traffic demand in 5G.

Applications and services need to be built with smart information so that M2M devices can sleep for as long as possible, only waking up when the information they possess is actually needed.

RESILIENCY AND ROBUSTNESS

The levels of resilience and robustness required by mission-critical services tend to be much higher than, for example, subscriber services, so 5G systems must be able to deliver these levels. In practical terms, one of the most significant challenges will be to provide support for mission-critical business services, as well as some legacy services.

High-grade network slices, which are capable of providing the desired levels of availability, robustness and resilience to attacks, can be offered to emergency services – with the additional option for extra capacity or coverage on demand, in the event of an emergency.

USE CASES

MOBILE BROADBAND

Existing mobile-broadband technologies like LTE will continue to evolve their capabilities to provide the backbone of the overall radio-access solution beyond 2020. One such capability might be the provision of data rates of several 100Mbps for use in urban and suburban environments and rates of several 10Mbps essentially everywhere – including sparsely-populated rural areas in both developed and developing countries. Smart antennas, including a very large number of steerable antenna elements, more spectrum and coordination among base stations will help to provide these levels of service to 5G users.

MEDIA

Ericsson's vision for the Networked Society includes a prediction for 15 billion video-enabled devices by 2020 [4]. To a large extent, these will be industrial video devices for machine monitoring, remote medicine, security, surveillance and image recognition, and large amounts of media data will be carried on both the uplink and the downlink.

The 5G systems will need to support near-zero latency interactions for use cases like remote medicine, as well as more cost- and radio-resource efficient operations with relaxed latency requirements.

In particular, uplink video will become more significant than it is today, as professional and user-generated content will be uploaded from cellular systems. Devices like cameras and wearables will be equipped with 5G transmitters, enabling continuous streaming of content in the uplink to, for example, a cloud-based production site to optimize production costs.

Video will be an important tool for emergency services like the police, fire brigade, and ambulance to provide civil protection and perform emergency management. Efficient media delivery with ultra-low latency for group call services, or relaxed latency for distribution of warning messages, will provide all teams involved in, for example, a rescue operation with access to the same media information, such as video streams from helmet-mounted or helicopter-mounted cameras. This use case requires both uplink and downlink network capacity with very high availability and high speed mobility requirements. However, such features will be provided in a more cost- and resource-efficient manner, so that resource intensive operation modes are only used when needed.

MACHINE-TYPE COMMUNICATION

The requirements of machine-type communication (MTC) vary considerably. At one end of the scale lies massive MTC, with critical MTC services at the other end.

Massive MTC

Examples that fall into this use-case category include the monitoring and automation of buildings and infrastructure, smart agriculture, logistics, tracking and fleet management. The requirements for massive MTC include:

- architecturally simple devices that use a low-complexity transmission mode
- devices that can run on battery power for many years
- long transmission ranges for devices in remote locations
- scalable networks that can connect either a large or a small number of M2M devices.

LTE already includes functionality to address the specific needs of MTC, and further improvements are envisaged for 5G radio access, including:

- support of different transmission modes such as device-to-device communication
- flexible sharing of radio resources between MTC services and mobile broadband services
- relaxed requirements on MTC devices, such as data rates, limited bandwidth, limited peak rates and half-duplex operation.

Network slices that support the MTC need for scalability can be optimized and support the introduction of leaner procedures such as reduced signaling – which is a significant aim, as signaling is the dominant part of the overall traffic (and consequently battery consumption) compared with the actual data transferred by massive MTC devices.

Critical MTC

In this type of communication, monitoring and control occur in real-time, E2E latency requirements are very low – at millisecond levels – and the need for reliability is high. The automation of energy distribution in a smart grid exemplifies the critical MTC case. In this case, the energy sources are volatile and distributed. The grid needs to manage the dynamics of volatile energy supply with demand, to protect the overall grid against faults that can occur. The radio access provided by 5G systems can guarantee low latencies, and the network slice can be configured so that network and application functions are physically placed in the network to ensure low latency, reliability and redundancy E2E.

Many different MTC use cases with varying requirements for connectivity exist in between the two ends of the MTC scale. Two representative use cases relate to remote machinery and intelligent transportation systems (ITSs).

Remote machinery

Future 5G systems will need to support remote control of heavy machinery – like excavators in mines and wood processors in forests. The benefit of this type of use case is to remove the need for people to work in hazardous environments or maybe to increase efficiency – through the ability to manage several machines simultaneously – or time can be saved when drivers are no longer required on location.

Crucial to this use case is the ability to transport high quality video, audio and other sensory information from the remote machine's environment to the controller. To control machinery in real time, the latency of the communication link between the machine and the controller needs to be extremely low. This necessity not only puts requirements on radio access, but also on the transport and core networks. In such scenarios, long transport links should be avoided, and processing may need to be moved closer to the machinery or to the remote controller. In addition, as the machinery or driver are likely to be on the move, the supporting network will need to be able to adapt to mobility.

ITSs

As modern cities become megacities and the migration from rural to urban environments continues, the need for good commuter systems will be significant to stem the rise of health/stress related problems, meet carbon emission goals, improve efficiency and integrate other industries with the transport system.

A fully interconnected transport system that includes all aspects of travel, including air, rail and bus, and commuter services linked to automobiles, people and bikes will provide a holistic view of any town, city, village or region.

For ITSs, information may be provided by fixed base stations alongside roads, railways and subway stations or from vehicles, people and trains. The challenge is to create a network slice with sufficiently low latency to provide useful alerts relating to approaching hazards.

CONCLUSION

With connectivity at the heart of industry transformation, 5G systems have a significant role to play – not just in the evolution of communication but in the evolution of businesses and society as a whole. Network slices are key to delivering differentiated offerings, as they can provide connectivity that is adapted and optimized for each and every use case, application and user – and do this in such a way that uses network resources efficiently.

As a natural evolution of current network architecture, broken apart into building blocks through SDN, NFV and virtualization technologies, 5G systems will provide a greater level of abstraction. Network functions can then be chained together as logical network slices supported by management and orchestration along with a flexible radio access environment.

REFERENCES

[1] Ericsson, June 2014, Ericsson Mobility Report, available at:
http://www.ericsson.com/res/docs/2014/ericsson-mobility-report-june-2014.pdf
[2] Ericsson Review, February 2014, Radio network energy performance: shifting focus from power to precision, available at: http://www.ericsson.com/news/140218-er-radio-network-energy-performance_244099438_c?
[3] EARTH, July 2012, Deliverable 6.4: Final Integrated Concept, available at:
http://bscw.ict-earth.eu/pub/bscw.cgi/d49431/EARTH_WP6_D6.4.pdf
[4] Ericsson, April 2014, Ericsson Media Vision 2020 – maps out the future of TV, available at:
http://www.ericsson.com/news/1775026

FURTHER READING

- Ericsson, October 2014, 5G: What is it?, available at:
  http://www.ericsson.com/res/docs/2014/5g-what-is-it.pdf
- Ericsson, October 2014, 5G: What is it for?, available at:
  http://www.ericsson.com/res/docs/2014/5g-what-is-it-for.pdf
- Ericsson Review, June 2014, 5G radio access, available at:
  http://www.ericsson.com/news/140618-5g-radio-access_244099437_c
- NGNM Alliance, December 2014, 5G White Paper – Executive Version, available at:
  http://www.ngmn.org/uploads/media/141222_NGMN-Executive_Version_of_the_5G_White_Paper_v1_0_01.pdf

GLOSSARY

CEM	customer experience management
CRM	customer relationship management
DC	data center
E2E	end-to-end
FCN	function
ITS	intelligent transportation system
LAA	Licensed Assisted Access
LSA	Licensed Shared Access
M2M	machine-to-machine
mmW	millimeter wave
MTC	machine-type communication
NF	network functions
NFV	Network Functions Virtualization
PKI	public key infrastructure
RAT	radio-access technology
RoT	Root of Trust
SDN	software-defined networking
SLA	Service Level Agreement
SSLA	Security Service Level Agreement
UE	user equipment
UI	user interface
vCSCF	virtual Call Session Control Function
vEPC	virtual Evolved Packet Core
vEPG	virtual Evolved Packet Gateway
vIMS	virtual IP Multimedia System
VM	virtual machine
vMSP	virtual multi-sequence positioning
VNF	virtualized network function