The Evolution Of Cellular IoT

A Technical Paper Prepared for SCTE•ISBE by

Hani Beshara
Diretor, MANA Radio Network Solutions
Ericsson Inc.
6300 Legacy Drive, Plano, Texas 75024
(214) 906-8231
hani.beshara@ericsson.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

Introduction

The IoT market, from networks to devices to use cases and applications, is evolving at a record speed. It is set to unleash a major transformation that has not been seen since the industrial revolution. IoT promises to change how we live and interact with the world around us, and to transform business productivity measures. It will help to finally unwire the factory and to provide the next level of automation such that industries can operate more efficiently and offer optimal products & services.

Cellular IoT is not a new concept. It is widely adopted across the globe, with 2G and 3G networks providing Low Powered Wide Area Networks (LPWAN) connectivity enabling many early IoT applications. As per Ericsson's mobility report, approximately 400 million 2G cellular connected devices have been in operation since early 2016. 4G LTE helped provide greater bandwidth, lower latency and increased support for large volumes of IoT devices per cell. By the end of 2018, the number of connections reached 1 Billion cellular connection. These will be enhanced further with the arrival of 5G networks, initially enabled by the 5G New Radio (NR) standard, which will enable Ultra-Reliable Low Latency Communications (URLLC) that support increasingly critical applications.

The remarkable cellular IoT growth rate is expected to continue, and the number of devices connected by Massive IoT and other emerging cellular technologies is forecast to reach 4.1 billion by 2024.

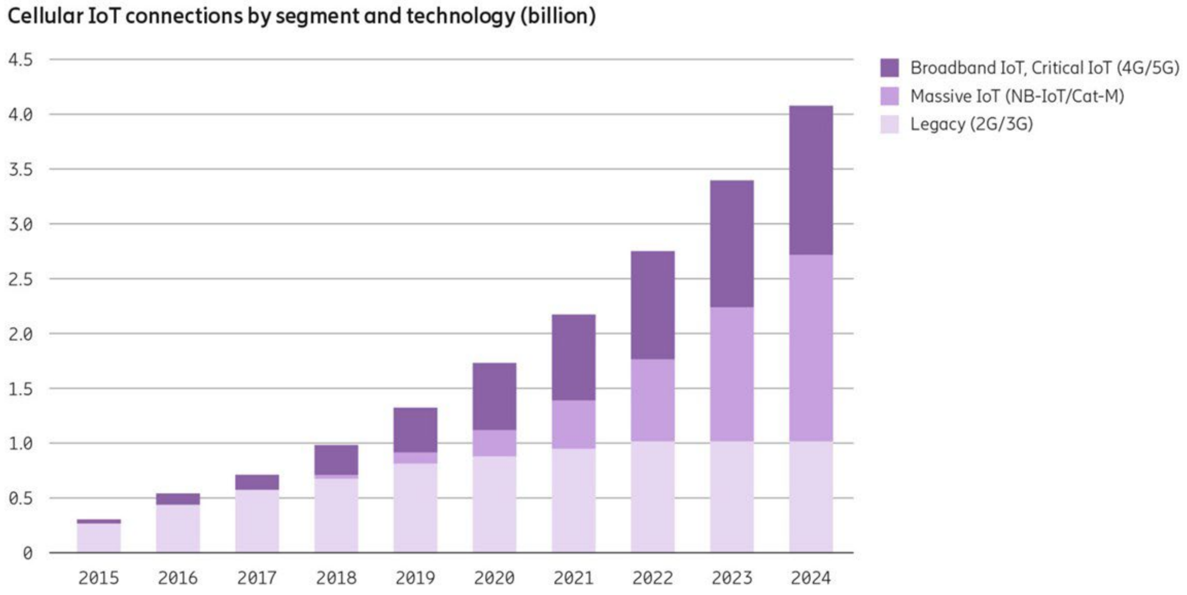


Figure 1 Ericsson Mobility Report – June 2019

Cellular IoT is supported by a rapidly growing ecosystem based on 3GPP global standards, offered by an increasing number of mobile network providers, and developed by a large number of device, chipset, module, and network infrastructure vendors. It offers unmatched global coverage in virtually every country in the world, Quality of Service, scalability, security and the flexibility to handle a varying set of requirements for a comprehensive range of use cases.

Four cellular IoT market segments are emerging with different requirements. These are:

- Massive IoT: low cost devices, small data volumes, massive numbers
- Broadband IoT: high throughput, low latency, larger data volumes
- Critical IoT: ultra-reliability, ultra-low latency, very high availability
- Industrial IoT: industrial protocols, time sensitive networks, precise indoor positioning

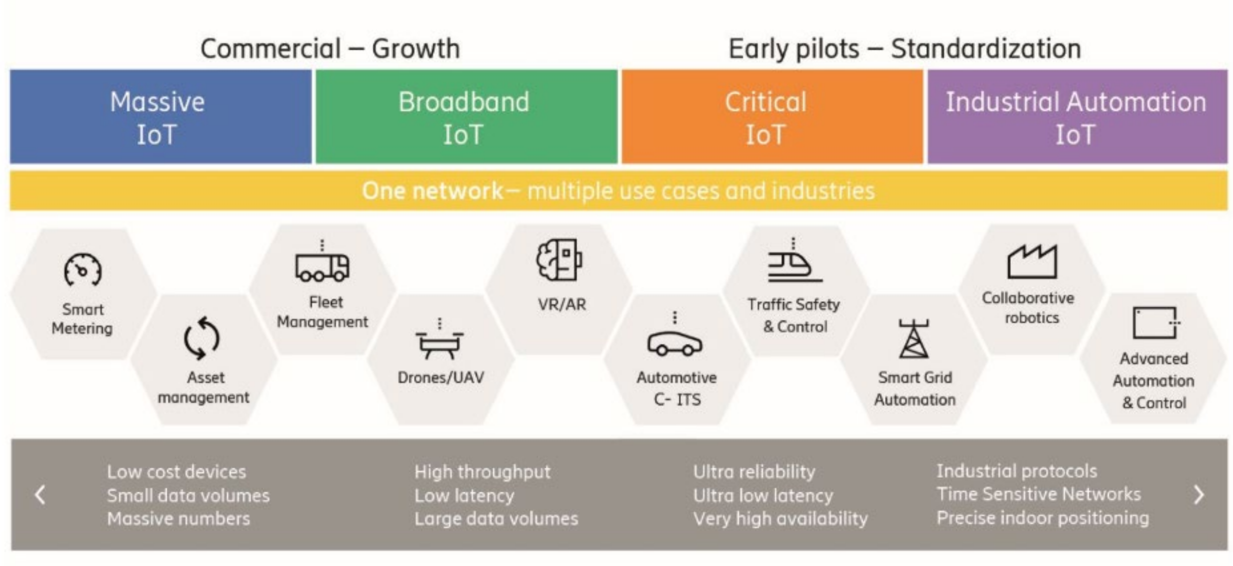


Figure 2 Cellular IoT Segments

This paper will review the emerging segments and the current and future support for the specific requirements within those segments through LTE and NR wireless technologies.

Content

1. Overview of cellular IoT market segments

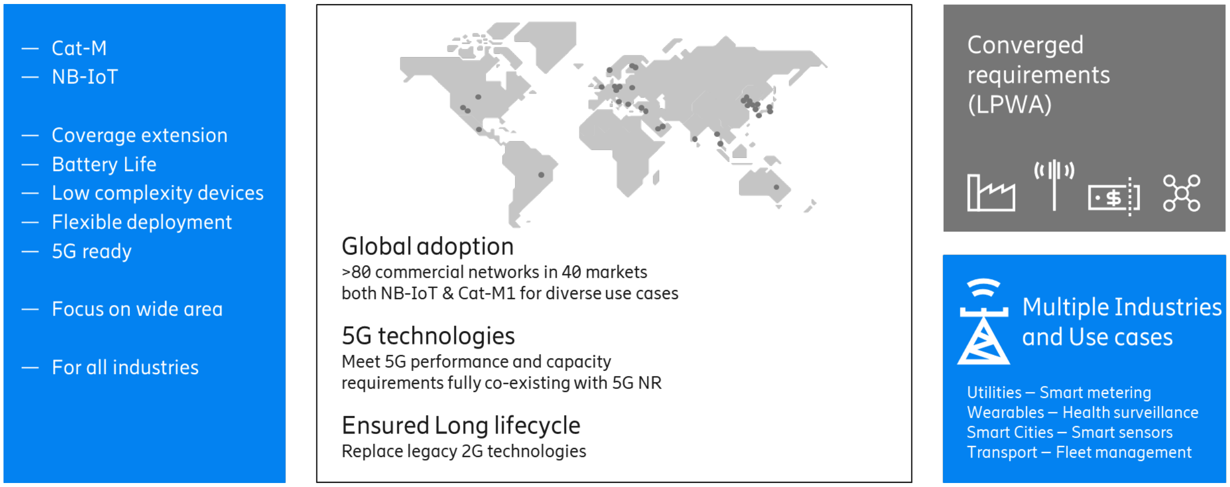
Massive IoT is a commercial cellular IoT market segment that is enabled by the 3GPP Release 13 standards which introduced the low complexity IoT variants; LTE-M and NB-IoT. As of early 2019, there are 80+ massive IoT networks in operation around the world. The Massive IoT segment has been growing rapidly with a compounded annual growth of 27 percent expected between 2018 and 2024 (Ericsson Mobility Report).

The Broadband IoT cellular segment refers to the existing IoT use cases that are enabled using 4G LTE, as well as the future use cases that would require 5G NR high throughput networks.

Critical IoT and Industrial IoT are emerging segments that will be made possible with enhancements associated with 5G NR and 3GPP releases 14/15/16+ by enabling Ultra reliable low latency communications (URLLC) and industrial protocols used for manufacturing.

1.1. Massive IoT segment

Massive IoT connectivity targets huge volume of low-complexity devices that infrequently send or receive small messages. The traffic is tolerant of delay and typical use cases include low-cost sensors, meters, and trackers. Such devices are often deployed in challenging radio conditions such as in basement of a building. Therefore, they require extended coverage and may rely solely on a battery power supply which puts extreme requirements on the device's battery life.



- Cat-M
- NB-IoT

- Coverage extension
- Battery Life
- Low complexity devices
- Flexible deployment
- 5G ready

- Focus on wide area

- For all industries

Global adoption
>80 commercial networks in 40 markets
both NB-IoT & Cat-M1 for diverse use cases

5G technologies
Meet 5G performance and capacity
requirements fully co-existing with 5G NR

Ensured Long lifecycle
Replace legacy 2G technologies

Converged requirements (LPWA)

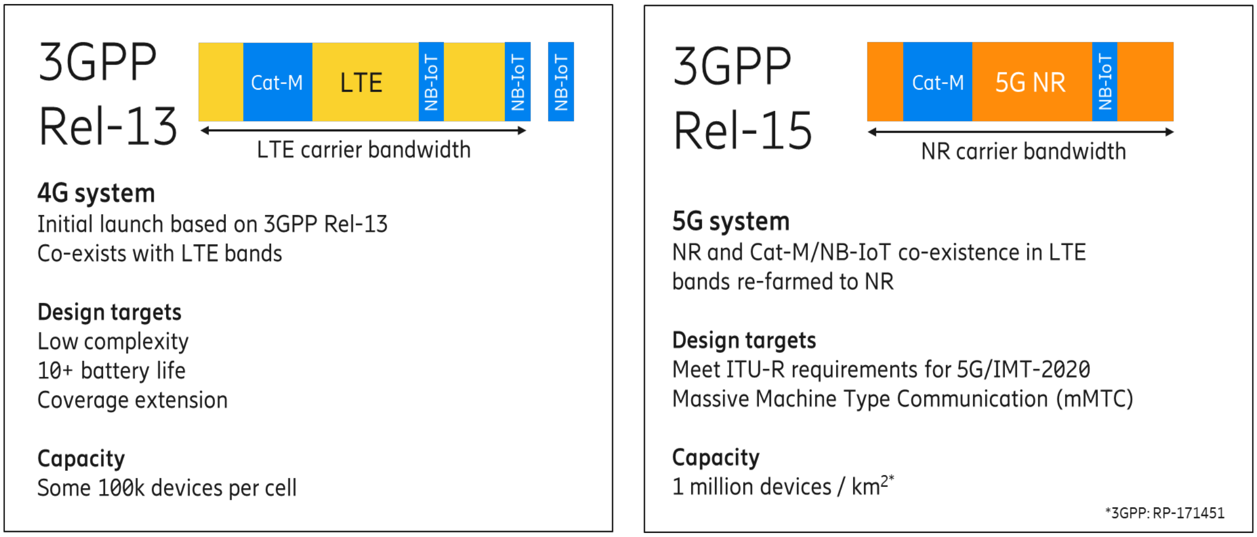
Multiple Industries and Use cases

Utilities – Smart metering
Wearables – Health surveillance
Smart Cities – Smart sensors
Transport – Fleet management

To support these attributes, new technology variants have been developed in the 3GPP standards Release 13; LTE-M and NB-IoT. Of the two, LTE-M (or CAT-M) supports greater bandwidth and complexity and is suited for wearables, trackers, alarm panels and customer support buttons, and hundreds of other use cases. LTE-M can support a VoLTE voice connection in addition to data, and supports bandwidth up to 1.4 Mhz.

NB-IOT is a stand-alone radio access technology that is designed for ultra-low complexity devices. NB-IoT only devices cost below $5 and support a 180 KHz bandwidth (similar to GSM technology). NB-IoT does not support voice, but supports superior coverage extension, up to 100 Kms. As GSM networks are sunset, it is expected that its IoT traffic will migrate to NB-IoT.

LTE-M and NB-IoT are designed as future proof technologies that will coexist with 5G network as 4G cellular networks, either evolve to, or are totally replaced by 5G networks. Both massive IoT technologies are being added to 3GPP Releases 15/16.



3GPP Rel-13

4G system
Initial launch based on 3GPP Rel-13
Co-exists with LTE bands

Design targets
Low complexity
10+ battery life
Coverage extension

Capacity
Some 100k devices per cell

3GPP Rel-15

5G system
NR and Cat-M/NB-IoT co-existence in LTE bands re-farmed to NR

Design targets
Meet ITU-R requirements for 5G/IMT-2020
Massive Machine Type Communication (mMTC)

Capacity
1 million devices / km^{2*}

*3GPP: RP-171451

3GPP IoT technology uses licensed spectrum comprising of pre-existing LTE bands which provides superior capacity and channel management, lower inter-device interference, reliability, and QoS. 3GPP service providers can provide scalable IoT services to their customer to meet their IoT device requirements in any geographical area.

Also in the context of LPWAN, other competing technolgoies, such as SigFox and LoRa, provide data service for IoT devices using the unlicensed spectrum. Such technologies enable new low cost networks and devices, however, they could suffer from added interference from other technologies using the same frequency resulting in limited capacity, limited bandwidth, and is usually under strict government regulations for both Tx power and the duty cycle. In Europe, SigFox and Lora regulations limit device transmission power to 25mW (14dBm) and the duty cycle of channel access to 0.1 percent or 1 percent of channel time. The unlicensed technology is expected to be used for the low cost, low bandwidth, and low reliability use cases. For a detailed comparison between licensed cellular IoT technolgoies and unlicensed IoT options, please refer to 5G America's white papar "LTE and 5G Technolgoies Enabling the Internet of Things".

1.2. Broadband IoT segment

Broadband IoT connectivity use cases require superior performance with low latency and high throughput. Typical applications include advanced wearables, aerial and ground vehicles, AR/VR enabled devices and sensors that require greater capabilities than what CAT-M or NB-IoT can provide. LTE has a range of device categories well-suited for such applications including a lower cost CAT-0 device. For example, LTE is already providing cellular connectivity to millions of modern cars for infotainment as well as preventive maintenance. There are LTE capable smart watches in the market today and LTE-connected drones are in the proof of concept stage. In support of this IoT market segment, LTE offers high spectral efficiency and data rates, low latencies and has been enhanced with extended device battery life and improved coverage. With advanced multi-antenna solutions and carrier aggregation, LTE enables peak rates in excess of 1 Gbps. Added to this, there are mechanisms for fast connection establishment and data delivery. With instant transmission schemes, the radio interface latency can be as low as 10ms. LTE scheduler can also support advanced priority handling mechanisms to provide superior performance to a selected group of users.

With 5G NR, the segment will expand with even greater bandwidth availability and lower latency of ~5ms over the air that can be suited for AR/VR and connected vehicle type applications.

- LTE
- 5G NR

- Multi-Gigabit
- Reduced latency
- Large data volume

- Focus on wide area

- For Automotive, Drones, Manufacturing, Utilities

Cooperative Intelligent Transport System (C-ITS) – V2X In Car MBB Services

Utilities – Smart Grids
Connecting grid elements beyond meters

Train Networks
Public Safety Networks

Drones – UAVs
Delivery, Infrastructure inspection, Agriculture

Advanced wearables
Health monitors, smart watches

1.3. Critical IoT segment

Ultra Reliable Low Latency Communications (URLLC) is a particular focus of the 5G NR standards in releases 15, 16, and 17 in support of the critical IoT market segment. Critical IoT use cases require extreme reliability with down time of (10^{-5} to 10^{-9}) mins per year and low latency (1ms and lower) over the air. Examples include smart grid control, fault restoration, real time control of machinery, intelligent transportation systems, remote surgery, and fully immersive AR/VR.

Early releases of 5G NR reduce latency and increases reliability over LTE. NR spectrum with flexible numerology supports 15 KHz to 120 KHz subcarrier spacing which will in turn reduce frame alignment delay and help reduce NR latency up to 1/8 msecs. Also new features, such as, Instantaneous DL assignment & UL dynamic scheduling, ultra-short duration transmission have been standardized to achieve the ultra low latency requirement. From redundancy perspective, 3GPP releases 16 and 17 are adding features for increased reliability including multi-connectivity, diversity, Robust coding and modulation, rapid retransmission protocols, prioritization mechanisms, and multiple signal transmission formats to achieve a reliability goal for 99.99999% and higher.

Critical IoT segment is expected to also take advantage of end to end network slicing and 5G NR QoS template to provide industries and private networks with its performance guarantees.

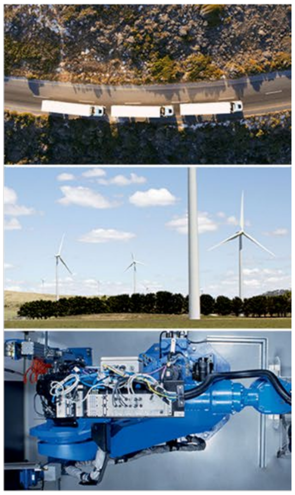
While wireless network price per bit has been on the decline, URLLC use cases is expected to have a much higher cost and price level.

- 5G NR
- URLLC
- 99.999% reliability
- 1ms one-way
- Local area/Wide area
- Automotive Utilities, Smart Manufacturing in industry campus

- Autonomous vehicles
- Deeper integration with C-ITS systems
- Platooning

- Utilities – Smart Grids
- Renewables integrated into Grid
- Real time control

- Real time control of industrial systems
- Fully immersive AR/VR



1.4. Industrial Automation IoT segment

Industrial Automation IoT builds on the capabilities of the three previous segments to address the specific needs of manufacturing and control systems for railways, power generation and power distribution.

This segment tailors the most demanding requirements from Manufacturing and Industrial Campuses and will support deterministic (and time sensitive) networks, industrial protocols such as PROFINET, EtherCAT, POWERLINK, and Modbus TCP/IP natively running over ethernet, together with very precise positioning.

The functionality and definitions for this segment are currently being defined in 3GPP, heavily influenced by Industry 4.0, and key industrial players and bodies such as 5G ACIA. It will be a 5G specific segment valid for local area and, almost certainly, non-public network deployments.

- 5G NR
- eURLLC
- Deterministic networks
- Ethernet support
- Time sensitive networking
- Local Area / Non-public networks
- Smart industries

- Smart Manufacturing, Railways, Power generation and distribution
- Automation for Robotics, Control systems and Process optimization

- Native 'Ethernet over NR' to support industrial protocols
- TSN and QoS
- Precise positioning

- 5G enabling Industry 4.0
- 5G-ACIA and other industry body collaborations

The ambition of this segment is to take that final step in the digitalization of the factory and move from today's mix of wired and wireless connectivity using multiple technologies into a wireless 5G network which will consolidate all requirements.

To fully support this segment, the standard must merge 5G NR and industrial control domains. 3GPP standards work is ongoing to enable precise indoor positioning, native support for Ethernet over NR, scheduling and QoS adaptations that will enable transparent Time Sensitive Networking (TSN) and allow reuse of existing industrial devices and control systems. With these additional capabilities, 5G NR have the potential to become the one common network that can support all existing use cases as well as fill the unfulfilled gaps in industrial communication and control networks.

2. IoT Use cases – additional technology considerations

In addition to the IoT connectivity aspects discussed above, there are few additional key considerations that are important for the success of all IoT use cases in all segments and for managing IoT devices over their complete life cycle; Security, Identity and Machine Intelligence.

2.1. Security

While securing a low price and a low complexity sensor may not seem as a critical activity, it has been proven that it can act as a trojan horse that opens a gate to the underlying network. Hence, IoT security needs to be taken seriously, even in what may seem as a simple application. Generally, Security of an IoT use case is a function of using secure communication, application security, and device security. Together, these functions protect the device, the network, guarantee data ownership, and ensure that trusted devices remain secure throughout their entire operational life.

Communication protocols like TLS, DTLS, and OSCORE are meant for IoT devices communications, however, not all supported algorithms are equally secure. Low complexity devices also have limited memory and processing power, hence it is even more important to select the optimal communication protocol. Newer protocols like TLS v1.3 are more secure and in many cases, are also more efficient.

In addition to secure communications, cryptographic keys and algorithms play an important role in securing IoT devices. IoT devices often only support the simpler symmetric key cryptographic algorithms (vs. public key cryptographic functions). However, with proper design (such as IETF Authentication and Authorization for Constrained Environments/OSCORE), it is possible to use public-key cryptographic functions in small IoT devices. The power consumption of complex computations can be reduced by using optimized hardware acceleration of cryptographic functions. It is therefore likely that future small IoT devices will have certain dedicated cryptographic hardware.

Persistent cryptographic keys must be stored securely and kept isolated from application software and physical interfaces. IoT devices can use the “Isolation” mechanism of Trusted Execution Environments (TEEs) to achieve this objective. Recently, ARM’s “TrustZone” TEE technology was brought to constrained devices. For more powerful devices, there are alternatives such as Intel SGX. Also, dedicated security components like Trusted Platform Modules or proprietary ASICs (application-specific integrated circuits) can be used. The goal is to develop solutions that can achieve a high level of security, albeit at higher cost and power consumption levels and in many use cases, integrated TEEs will be sufficient and more cost-effective.

To maintain security during their operational life, IoT devices should support secure software/firmware upgrade though “root of trust” mechanism. Secure upgrades are realized by having the software signed prior to release and having a trusted subsystem in the device perform a verification of the software before it is programmed/loaded into the device. New standardization work for securing updates for software/firmware. Procedures for secure device life-cycle management have been initiated. It is however complex and may have to be tailored for each specific use case.

The criticality of device security is growing in the industry, however, much more efforts are needed to develop trusted systems that cover the needs of life-cycle management and applications security. Supporting secure software update is crucial to the creation of trustworthy IoT devices.

2.2. Identity

A digital identity is crucial in identifying device trustworthiness and in the overall network security. By trustworthiness, we mean that a device can prove that is been produced by a legitimate manufacturer through an initial identity. A digital identity can be used for authentication, to maintain data ownership, or for software origin verification.

An identity consists of a securely stored secret and an assigned link between this secret and an identifier or name. This can be accomplished using a public key infrastructure (PKI), where the device holds a private key and the identity is a certificate that links this key to an identifier written into the certificate. For IoT devices, traditional PKIs have their problems. Their cryptographic operations can be cumbersome for highly constrained low complexity devices, the certificates can be large, and the certificate revocation management is usually so complex and error prone that it is hardly used. Furthermore, traditional PKIs have privacy issues. These issues can be addressed, using Enhanced Privacy ID, but at significantly higher complexity costs.

As an alternative to PKIs, it is possible to use identities based on symmetric key cryptography. This method is already in use for the 2G, 3G and 4G mobile network systems that use SIMs to hold the authentication credentials. SIMs use dedicated hardware chips and are relatively complex, mainly for legacy reasons. More cost-effective solutions are on their way, such as the integrated Universal Integrated Circuit Card (iUICC), in which the SIM hardware is integrated into the device processors. For 5G mobile network systems, symmetric key based identities for network access will remain in use, but in 5G it is also possible to use PKI-based identities via Extensible Authentication Protocol (EAP)-TLS. Figure below illustrates EAP-TLS ID management and use for network access.

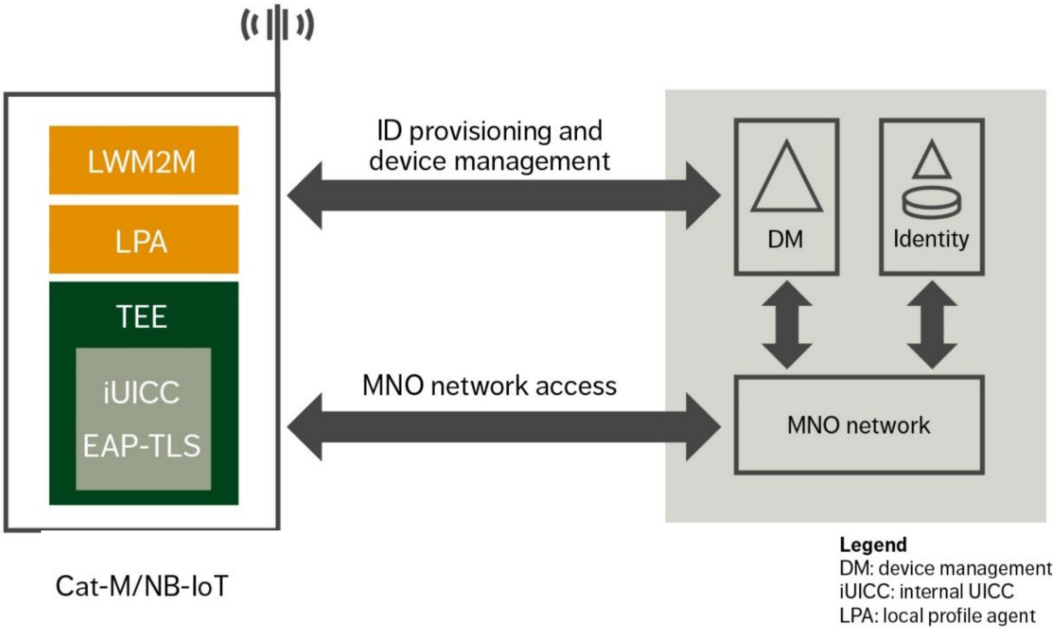


Figure 3 EAP-TLS ID management

Blockchains can play even a more central role in the distributed approach to handling the trust in device identities. These options make it possible to link device lifecycle management with that of the device identity in a common framework.

2.3. Machine Intelligence

MI technologies are essential to building IoT systems that can improve their own performance as more data becomes available and more knowledge is inferred and retained. With IoT and its associated large volume of data and billions of devices, MI is required to intelligently automate data collection and processing. Distributed MI (DMI) concerns the deployment, dynamic composition and life-cycle management of multi-node MI services, which can be chained for provisioning an intelligent system.

With distributed MI, it is possible to move the intelligence toward the device end (or the edge), which will minimize E2E latency, enhance data privacy and lower bandwidth requirements. Such on-device MI (ODMI) enables horizontal connectivity of devices to edge infrastructure that hosts DMI services. To accomplish this goal, IoT devices should be capable of performing low-power computation at or close to where the data is generated, or where the resulting action is needed. From a software perspective, it is possible to offload MI computation to hardware accelerators at the edge. In this layer, compilers and schedulers break down MI workflows, and distribute it into smaller tasks, which can help optimize the overall process.

Investment in scalable and flexible MI IoT systems will play a key role in evolving networks, as new devices, sensors, and actuators are being added and removed. Edge compute and ODMI are hence important in providing flexible deployment option and for managing new and complex Service Level Agreements.

Conclusion

IoT use cases and technology are evolving rapidly. 3GPP standards provide a holistic, optimal, and flexible solutions covering a wide range of use cases. 3GPP's LTE-M and NB-IoT are optimized to serve a variety of use cases with global reach where small amount of data are sent infrequently, with high latency tolerance, low cost, extended coverage, and low battery consumption.

Today's 4G LTE broadband networks serve a large number of existing IoT use cases that require low latency and/or high throughput (e.g. Vehicle OEM communication, security cameras, etc.). 5G NR and 3GPP releases 14/15/16+ will enable an emerging set of use cases in critical IoT that require ultra reliable low latency communications, as well as the support of industrial IoT protocols that are used in the manufacturing sector.

When considering an IoT solution, careful consideration of security, device identity and IoT data management are key to success. With regards to security, the implementation of cryptographic functions on the device is the optimal approach to achieving strong device security. TEEs will soon be applied to IoT devices to support use cases in which secure storage and isolation are required. Secure identities are important to identify the origin of data and to ensure secure connectivity. New cost-efficient solutions for LPWAN access will emerge, leveraging the device's built-in security capabilities. Machine Intelligence will make it possible to move processing toward the device end and the mobile which will minimize E2E latency, enhance data privacy, and lower bandwidth requirements.

Abbreviations

3GPP	Third Generation Partnership Project
AP	access point
ASIC	Application Specific Integrated Circuit
bps	bits per second
EAP	Extensible Authentication Protocol
EtherCat	Ethernet for Control Automation Technology
FEC	forward error correction
HFC	hybrid fiber-coax
HD	high definition
Hz	Hertz
iUICC	Integrated Universal Integrated Circuit Card
IoT	Internet of Things
ISBE	International Society of Broadband Experts
LPWAN	Low Power Wide Area Networks
LTE	Long Term Evolution
MI	Machine Intelligence
ML	Machine Learning
NR	New Radio – 3GPP 5G Radio Protocols
OEM	Original Equipment Manufacturer
OSCORE	Object Security for Constrained RESTful Environments
PROFINET	Portmanteau for Process Field Net
PKI	Public Key Infra-structure
QoS	Quality of Service
SCTE	Society of Cable Telecommunications Engineers

SGX	Software Guard Extensions
TEE	Trusted Execution Environment
TLS	Transport Layer Security
TSN	Time Sensitive Networking
URLLC	Ultra Reliable Low Latency Communications

Bibliography & References

Ericsson white paper, January 2019, Cellular IoT Evolution for Industry Digitalization

Ericsson Technology Review, January 9, 2019, Key technology choices for optimal massive IoT devices

Ericsson Mobility Report – June 2019

5G America's white papar - LTE and 5G Technolgoies Enabling the Internet of Things