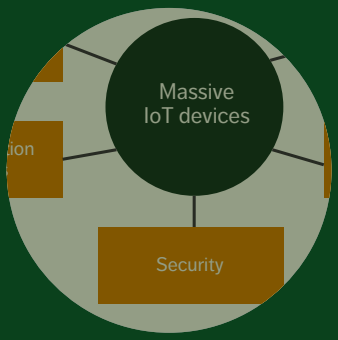
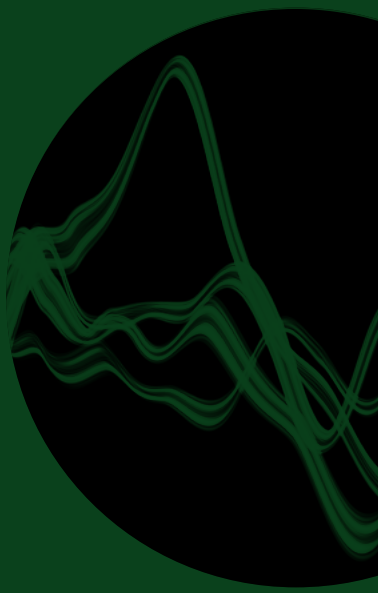
CHARTING THE FUTURE OF INNOVATION | #12·2018

ERICSSON TECHNOLOGY Review





MASSIVE IoT DEVICES

ERICSSON

KEY TECHNOLOGY CHOICES FOR

optimal massive IoT devices

The latest cellular communication technologies LTE-M and NB-IoT enable the introduction of a new generation of IoT devices that deliver on the promise of scalable, cost-effective massive IoT applications using LPWAN technology. However, a few key technology choices are necessary to create IoT devices that can support the multitude of existing and emerging massive IoT use cases.

CLAES LUNDQVIST,
ARI KERÄNEN,
BEN SMEETS,
JOHN FORNEHED,
CARLOS R. B. AZEVEDO,
PETER VON WRYCZA

The Internet of Things (IoT) represents an ongoing paradigm shift within communications: everything that benefits from a connection can and will be connected.

■ Massive IoT refers to applications that are less latency sensitive and have relatively low throughput requirements, but require a huge volume of low-cost, low-energy consumption devices on a network with excellent coverage. The growing popularity of IoT use cases in domains that rely on connectivity spanning large areas, and are able to handle a huge number of connections, is driving the demand for massive IoT technologies.

Through the development of new technologies in the fields of communication, computation, sensors, electronics and batteries, it is now possible to develop battery-powered devices with sensors and actuators and computers that are connected via wide-area communication networks to a cloud-based platform that handles device data and management. These devices can be tailored to fit several specific application areas and deployed in massive numbers, making them fit for use in massive IoT applications.

Examples of massive IoT application areas include: wearables (e-health); asset tracking (logistics); smart city/smart home, environmental monitoring and smart metering (smart building); and smart manufacturing (monitoring, tracking, digital twins). The key device characteristics include:

- low device and deployment cost
- small form factor
- long battery life
- wireless connectivity for challenging locations
- strong application and communication security.

There are two key challenges in the massive IoT device domain: (1) connecting a large volume

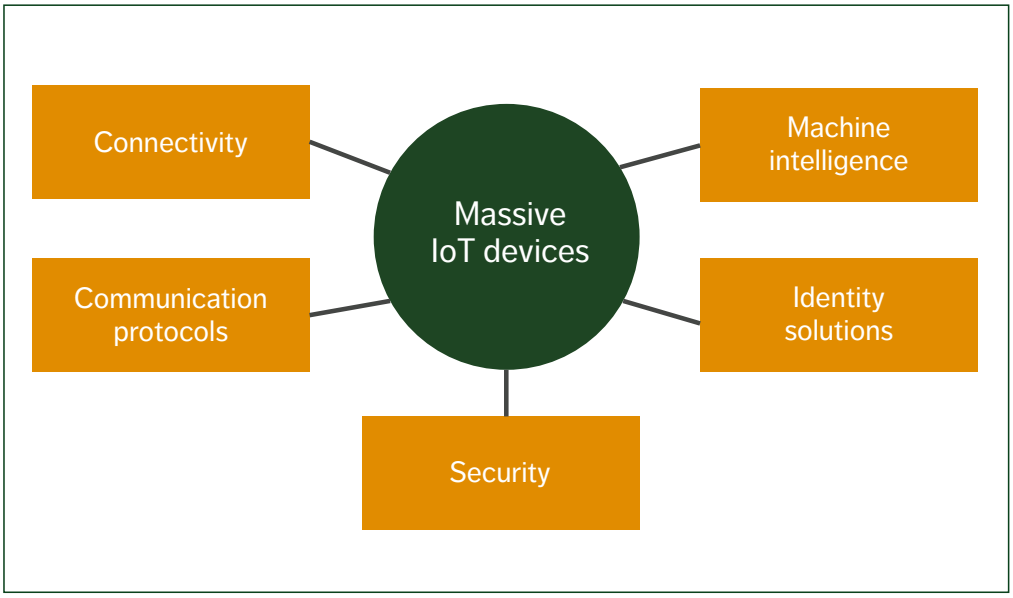


Figure 1 Key technologies for massive IoT devices

of devices in a wide area cost-efficiently, and (2) efficiently managing these devices over their complete life cycle. As security and trust are key requirements in most massive IoT applications, the devices must be trusted in terms of both communication and data integrity end-to-end (E2E), from device to application data usage. Many applications also benefit from devices that include local intelligence that can process data before it is further communicated.

To address these challenges, it is necessary to make smart choices in five key technology areas – connectivity, communication protocols, security, identity solutions and machine intelligence (MI) – as shown in *Figure 1*. Carefully considered choices in these five areas make it possible to achieve the desired key device characteristics and create IoT devices that support the multitude of existing and emerging massive IoT use cases.

Connectivity

New massive IoT cellular technologies, such as Narrowband IoT (NB-IoT) and LTE for machine-type communication (LTE-M), are taking off and driving growth in several cellular IoT connections, with a compound annual growth rate of 27 percent expected between 2018 and 2024 [1]. LTE-M and NB-IoT are cellular radio access technologies that provide low-power wide-area (LPWA) IoT connectivity in licensed spectrum, unlike short-range technologies in unlicensed spectrum such as Bluetooth and Zigbee, and LPWA technologies such as Sigfox and LoRaWAN.

The 3GPP release 13 design targets for massive IoT were: long device battery life, low device complexity to ensure low cost, support for massive numbers of devices, and coverage enhancements to be able to reach devices in basements and other challenging locations. Two new cellular technologies

were introduced in 3GPP release 13: LTE-MTC (LTE-M), which includes a new user equipment (UE) category called Cat-M1, and NB-IoT, which includes UE category Cat-NB1 [2].

A Cat-M1 UE supports a reduced bandwidth of 1.4MHz and a data throughput of up to 300kbps in the downlink and 375kbps in the uplink. It also supports mobility and VoLTE services. Therefore, Cat-M1 UEs are suitable for applications such as wearables and asset tracking.

NB-IoT operates in half-duplex mode within the 200kHz bandwidth and supports a data throughput of up to 26kbps in the downlink and 63kbps in the uplink. Similar to Cat-M1, NB-IoT offers the coverage enhancement feature, with up to +20dB enhanced coverage, versus +15dB in Cat-M. The UE output power classes are 20dBm and 23dBm, as in Cat-M.

To improve the user experience and to cater to more use cases, several enhancements and new functionalities are introduced in 3GPP LTE-M and NB-IoT releases 14 and 15 [3] [4]. Among other things, release 14 features improvements to LTE-M – such as more accurate positioning of UE, multicast transmission and VoLTE in enhanced coverage, as well as higher data rates to serve a wider range of applications, reduce latency and extend battery life.

Similarly, release 14 NB-IoT performance is improved with more accurate positioning of UE, multicast transmission, capacity improvement (thanks to the support of paging and random-access procedures on non-anchor carriers), higher peak data rates and a new lower power class (14dBm) that enables reduced power consumption and smaller battery form factors.

In release 15, LTE-M features include support for higher UE velocities, a new lower UE power class, reduced system acquisition time, reduced UE power consumption by early data transmission, a wake-up signal for paging monitoring, relaxed monitoring for cell reselection, increased spectral efficiency and improved access control.

The main features introduced in release 15 NB-IoT aim to further reduce latency and UE power consumption (early data transmission, wake-up signal and quick Radio Resource Control release, for example). Other features include: UE measurement improvements, support of cell ranges of up to 100km, TDD support, reduced system information acquisition and cell search time, and improved UE differentiation and access control.

Terms and abbreviations

ASIC – Application-Specific Integrated Circuit | **CoAP** – Constrained Application Protocol | **DMI** – Distributed Machine Intelligence | **E2E** – End-to-end | **EAP** – Extensible Authentication Protocol | **HTTP** – Hypertext Transfer Protocol | **IETF** – Internet Engineering Task Force | **IoT** – Internet of Things | **IPSO** – Internet Protocol for Smart Objects | **iUICC** – Integrated Universal Integrated Circuit Card | **LoRaWAN** – Long Range Wide-Area Network | **LPWA** – Low-Power Wide-Area | **LPWAN** – Low-Power Wide-Area Network | **LTE-M** – LTE for Machines | **LwM2M** – Lightweight M2M | **M2M** – Machine-to-Machine | **MI** – Machine Intelligence | **MNO** – Mobile Network Operator | **MQTT** – Message Queuing Telemetry Transport | **MTC** – Machine Type Communication | **NB-IoT** – Narrowband Internet of Things | **ODMI** – On-Device Machine Intelligence | **OSCORE** – Object Security for Constrained RESTful Environments | **PKI** – Public Key Infrastructure | **QUIC** – Quick UDP Internet Connections | **SenML** – Sensor Measurement Lists | **SGX** – Software Guard Extensions | **TCP** – Transmission Control Protocol | **TEE** – Trusted Execution Environment | **TLS** – Transport Layer Security | **TPU** – Tensor Processing Unit | **UDP** – User Datagram Protocol | **UE** – User Equipment | **WoT-TD** – Web of Things Thing Descriptions

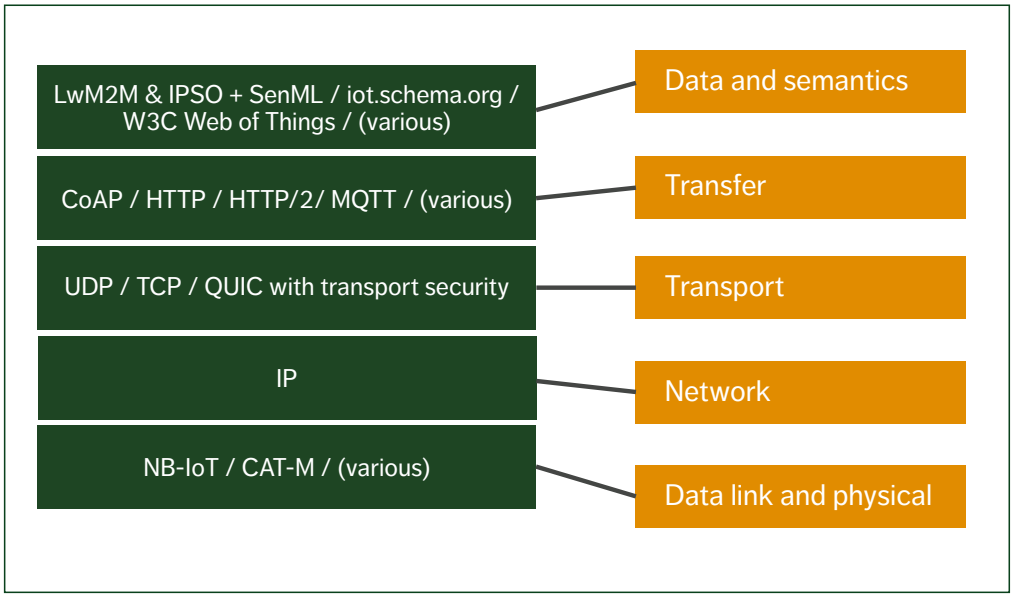


Figure 2 Structure of an IoT device protocol stack

Communication protocols

While many legacy machine-to-machine (M2M) devices use tailor-made protocol stacks for each specific application, more and more devices today (as well as the vast majority of current ecosystems) use internet protocols as the basis of the IoT protocol stack. That is, they use the Internet Protocol (IP) on top of various data link protocols, followed by a selection of standardized transport and transfer protocols, ending up at the application layer with data models and semantics, as shown in *Figure 2*.

The latest compression techniques, such as Static Context Header Compression [5] can compress the IPv6 and other headers into just a few bytes, making it possible for even the most constrained low-power wide-area network (LPWAN) IoT communication systems to use IPv6. On top of IPv6, User Datagram Protocol (UDP) or Transmission Control Protocol (TCP) is usually used at the transport layer. More recently, the QUIC protocol [6], combining features from UDP and TCP, is attracting interest for IoT scenarios as well.

IoT E2E communication is usually secured with Transport Layer Security (TLS). Recently, the Internet Engineering Task Force (IETF) finished the standardization of TLS v1.3. This latest version enables faster connection setup, more resiliency to address changes and stronger security. When E2E security through middleboxes, such as proxies, is needed, IoT communication can be secured with Object Security for Constrained RESTful Environments (OSCORE) [7].

Transfer protocols are used over the (secure) transport layer to transfer data objects and provide semantics for operations. Two transfer protocols that reuse the web model are widely used today: Hypertext Transfer Protocol (HTTP) [8] and Constrained Application Protocol (CoAP) [9]. The new version of HTTP, HTTP/2 [10], is also increasingly being adopted. Message Queuing

> IT IS POSSIBLE TO TAKE ADVANTAGE OF PUBLIC-KEY CRYPTOGRAPHIC FUNCTIONS IN SMALL IOT DEVICES

Telemetry Transport (MQTT) is a widely-used publish-subscribe protocol for the IoT. In industrial environments, more specialized protocols are often used, and some environments also reuse legacy messaging protocols for IoT. Out of all the options, web protocols, and in particular CoAP for the embedded web, have proven to be the best choice, especially for interoperability and scalability.

Data models provide common syntax, structure and semantics for the communicating endpoints. A data model can be something very simple – containing a single temperature value, for example – but most real-life systems require the exchange of more information. Traditionally, in many M2M systems this information has been encoded in application-specific ways, but in the IoT, where data is often exchanged with multiple types of loosely coordinated systems, common data models are needed to ensure endpoints understand the meaning of the data. Standardized data models such as Sensor Measurement Lists (SenML) [11] can be used to efficiently interchange batches as well as the time series of sensor and actuator data.

A fully built and operational IoT system also requires life-cycle management capabilities such as automated bootstrapping, configuration and firmware updates. The Open Mobile Alliance SpecWorks Lightweight Machine-to-Machine (LwM2M) device and data management protocol [12] is built on the standard web protocol stack, using IP, UDP/TCP, CoAP, TLS/OSCORE and SenML. Furthermore, IPSO smart objects can be used with LwM2M to enable reusable application semantics. LwM2M and IPSO smart objects provide a full suite to support life-cycle management and applications with interoperability from connectivity to application layer.

Finally, it is possible to bridge the gap between devices from different – and often uncoordinated – ecosystems by using common ways to express device interaction capabilities such as the World Wide Web Consortium's Web of Things Thing Descriptions (WoT-TD) [13], and common vocabularies for describing things, such as iot.schema.org.

Security

The security of IoT devices is built on functions for secure communication, application security and device security. Together, these functions protect device management, guarantee data ownership and ensure that devices remain trustworthy throughout their entire operational life. Secure communication protocols like TLS, DTLS and OSCORE allow for different algorithms. However, not all supported algorithms are secure – this is the case for TLS v1.2, for example. In addition, IoT devices normally only support a subset of algorithms, which makes it important to select the right ones. Newer protocols like TLS v1.3 are more secure and in many cases also more efficient.

IoT devices often only support symmetric key cryptographic algorithms, due to the fact that public-key cryptographic functions are complex and demand large key sizes, which may be problematic for very constrained devices. With proper design (as in IETF Authentication and Authorization for Constrained Environments/OSCORE), however, it is possible to take advantage of 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 key material must be stored securely and kept isolated from application software and physical interfaces as much as possible. IoT devices are increasingly following the smartphone approach of using Trusted Execution Environments (TEEs) for this isolation. 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. Such solutions can achieve a high level of security, albeit at higher cost and power consumption levels. 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. Such secure upgrade is often realized by having the software signed prior to release and having a trusted subsystem in the device that performs a verification of the software before it is programmed/loaded into the device. This trusted subsystem is often referred to as the root of trust of a device. New standardization work [14] was recently started for securing updates for software/firmware. Procedures for secure device life-cycle management are not easy and may have to be tailored for a specific use case. The awareness of the importance of device security is growing in the industry, but more efforts are needed to realize well-integrated trustworthy 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.

Identity solutions

Trustworthiness also depends on secure digital identities. A digital identity can be used for authentication, to maintain data ownership or for software origin verification. For example, a device can prove it is trustworthy – that is, it has been produced by a legitimate manufacturer – through an initial identity.

An identity consists of a securely stored secret and an assigned link between this secret and an identifier or name. A well-known way to do this is to use 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 devices, the certificates can be large, and the certificate revocation management is usually so tricky that it is hardly used. Furthermore, traditional PKIs have privacy issues. These issues can be addressed, as they have been in Enhanced Privacy ID, but at significantly higher complexity costs than PKI.

“ SUPPORTING SECURE SOFTWARE UPDATE IS CRUCIAL TO THE CREATION OF TRUSTWORTHY IoT DEVICES ”

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

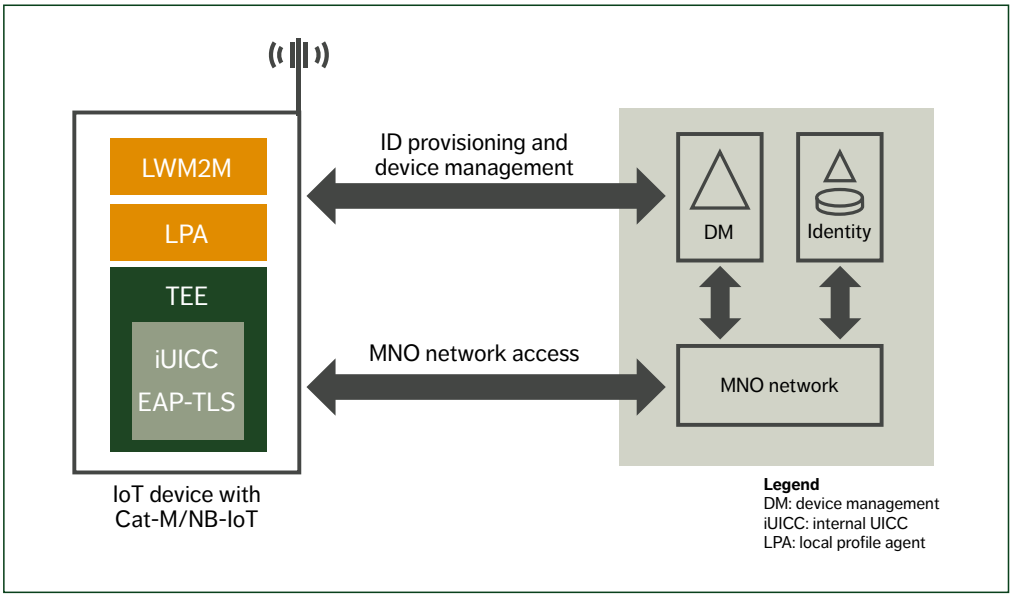


Figure 3 EAP-TLS ID management and use for network access

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 3* illustrates EAP-TLS ID management and use for network access.

Beyond mobile networks, other network technologies also require identities, and applications may need identities too. Therefore, depending on the device use case, a single device may need several identities. This can be problematic for constrained devices, and it makes identity management difficult. As different device hardware will come with different types of initial identities, Ericsson believes that a federation of identities [15] is important in the bootstrapping of identities that support the device use case.

The complexity of identity management can be reduced if identities can be reused. In practice, such reuse may be built on careful derivation techniques, in which a new identity is created and receives trust from an existing one. This is, for example, the case in Generic Bootstrapping Architecture, where a SIM-based key can be used to derive a key for TLS or application security.

A more holistic and distributed approach to handling the trust in device identities can be achieved with blockchains or distributed ledgers.

These options make it possible to link device life-cycle management with that of the device identity in a common framework.

Machine intelligence

MI technologies are key to building IoT systems that can improve their own performance of a task as more data becomes available and more knowledge is inferred and retained [16]. In massive IoT, which handles large volumes of data and millions of devices, MI is required to intelligently automate data transmission, routing and data 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. Orchestrating lightweight DMI components to jointly perform MI tasks that enhance massive IoT operations is a fundamental research topic at Ericsson [17].

One important path in DMI is moving intelligence toward the device end, which will minimize E2E latency, enhance data privacy and lower bandwidth requirements while reducing server-side costs. Such on-device MI (ODMI) efforts go beyond routing IoT data to cloud backends and instead promote horizontal connectivity of devices to edge infrastructure that hosts DMI services.

To follow this path, it is essential that the IoT devices are able to perform low-power computation close to where the data is generated and the actuation is needed. This requires knowledge of MI-tailored ASICs and of their integration with MI frameworks. In the hardware layer, ODMI has been embodied into graphics processing units, ASICs such as tensor processing units (TPUs), and neuromorphic chips. The main innovation of TPUs relies on efficient complex instruction set implementations for the matrix multiplier unit, which is key for executing modern MI workflows. Neuromorphic chips are low-power hardware where asynchronous brain-inspired manycore meshes are interconnected over sparse and recurrent inter-core communication topologies, thus easing the translation of MI dataflows into instruction flows.

On the software side, many vendors favor the idea of offloading MI computation to hardware accelerators. In this layer, the integration of systems optimization has become widespread, such as compilers and schedulers that can prune and break down MI workflows into distributable task graphs. Scalable massive IoT systems require investment in MI services that can be repurposed to adapt to operational conditions in evolving networks, as sensors and actuators are added and removed. Flexibility is then a core design principle in massive IoT systems. Edge and ODMI add such flexibility because they offer more DMI deployment options and control over changing Service Level Agreements.

> “ONE IMPORTANT PATH IN DMI IS MOVING INTELLIGENCE TOWARD THE DEVICE END”

Leading the MI and IoT convergence will require intertwining the right competence in unique team setups, bridging system architects, embedded systems designers and distributed system engineers, as well as subject matter experts on MI, security, IoT protocols and systems optimization. At Ericsson, we are taking this multidisciplinary challenge seriously to ensure that we are equipped to apply DMI competently to generate business value in emerging IoT markets.

Conclusion

Rapid technology advances in recent years have been of great benefit to the ongoing realization of massive IoT devices. It is, however, vital for device manufacturers, mobile network operators and other industry players to carefully consider the options and make the right choices when applying new technologies in the device domain. From Ericsson's perspective, there are five key technology areas that are of particular significance: connectivity, communication protocols, security, identity solutions and machine intelligence (MI).

In terms of connectivity, we are convinced that LTE-M and NB-IoT technologies will further enhance functionality and use-case applicability, improving the possibility to create devices with lower power consumption and a smaller form factor, at a lower cost. It is also our opinion that the best way to ensure the interoperability of IoT devices from communication to application layer is through the use of protocol stacks based on standardized internet protocols and data models with efficient capabilities for data transfer and device management.

With regard to security, we believe that 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 is crucial and isolation between functionality is required. It is also our view that the use of secure identities will soon become key, as a means to identify the origin of data and to realize secure connectivity. New cost-efficient solutions for LPWAN access will emerge, leveraging the device's built-in security capabilities.

Finally, advances in MI technologies have made it possible to move intelligence toward the device end, which we regard as a great opportunity to minimize E2E latency, enhance data privacy and lower bandwidth requirements, while reducing server-side costs.

Further reading

- **Ericsson web page, Internet of Things, available at:** *https://www.ericsson.com/en/internet-of-things*
- **Ericsson white paper, January 2016, Cellular networks for Massive IoT – enabling low power wide area applications, available at:** *https://www.ericsson.com/en/white-papers/cellular-networks-for-massive-iot--enabling-low-power-wide-area-applications*
- **Ericsson white paper, June 2017, IoT security – protecting the networked society, available at:** *https://www.ericsson.com/en/white-papers/iot-security-protecting-the-networked-society*
- **Ericsson white paper, March 2018, 5G security – enabling a trustworthy 5G system, available at:** *https://www.ericsson.com/en/white-papers/5g-security---enabling-a-trustworthy-5g-system*
- **Ericsson Research blog, March 2017, Smart contracts for identities, available at:** *https://www.ericsson.com/en/blog/2017/10/smart-contracts-for-identities*
- **Ericsson Technology Review, November 2017, End-to-end security management for the IoT, available at:** *https://www.ericsson.com/en/ericsson-technology-review/archive/2017/end-to-end-security-management-for-the-iot*

References

1. **Ericsson Mobility Report, November 2018, available at:** *https://www.ericsson.com/en/mobility-report/reports/november-2018*
2. **Academic Press, Cellular Internet of Things: Technologies, Standards and Performance, 1st edition, 2017, O. Liberg, M. Sundberg, E. Wang, J. Bergman, J. Sachs**
3. **IEEE Network, volume 31, issue 6, Overview of 3GPP Release 14 Enhanced NB-IoT, November/December 2017, A. Höglund et al.**
4. **IEEE Communications Standards Magazine, volume 2, issue 2, Overview of 3GPP Release 14 Further Enhanced MTC, June 2018, A. Höglund et al.**
5. **IETF, June 2018, LPWAN Static Context Header Compression (SCHC) and fragmentation for IPv6 and UDP, available at:** *https://tools.ietf.org/html/draft-ietf-lpwan-ipv6-static-context-hc-16*
6. **IETF, October 2018, QUIC: A UDP-Based Multiplexed and Secure Transport, available at:** *https://tools.ietf.org/html/draft-ietf-quic-transport-15*
7. **IETF, August 2018, Object Security for Constrained RESTful Environments (OSCORE), available at:** *https://tools.ietf.org/html/draft-ietf-core-object-security-15*
8. **IETF, June 2014, Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing, available at:** *https://tools.ietf.org/html/rfc7230*
9. **IETF, June 2014, The Constrained Application Protocol (CoAP), available at:** *https://tools.ietf.org/html/rfc7252*
10. **IETF, May 2015, Hypertext Transfer Protocol Version 2 (HTTP/2), available at:** *https://tools.ietf.org/html/rfc7540*
11. **IETF, August 2018, Sensor Measurement Lists (SenML), available at:** *https://tools.ietf.org/html/rfc8428*
12. **OMA SpecWorks, Lightweight M2M (LWM2M), available at:** *https://www.omaspecworks.org/what-is-oma-specworks/iot/lightweight-m2m-lwm2m/*
13. **W3C, October 21, 2018, Web of Things (WoT) Thing Description, available at:** *https://www.w3.org/TR/wot-thing-description/*
14. **IETF, Software Updates for Internet of Things (suit), available at:** *https://datatracker.ietf.org/wg/suit/about/*
15. **Intel, October 15, 2018, Intel and Arm Share IoT Vision to Securely Connect Any Device to Any Cloud, Lorie Wigle, available at:** *https://newsroom.intel.com/editorials/intel-arm-share-iot-vision-securely-connect-any-device-any-cloud/*
16. **Ericsson Technology Review, April 2017, Tackling IoT complexity with machine intelligence, available at:** *https://www.ericsson.com/en/ericsson-technology-review/archive/2017/tackling-iot-complexity-with-machine-intelligence*
17. **Ericsson white paper, May 2018, Artificial intelligence and machine learning in next-generation systems, available at:** *https://www.ericsson.com/en/white-papers/machine-intelligence*

THE AUTHORS

Claes Lundqvist

◆ serves as director of Technology Foresight at Ericsson Group Function Technology. He joined Ericsson in 1996 and has held various positions in R&D and product management, working with technology platforms for mobile devices. His current work focuses on the technology management area, including technologies for mobile devices and the IoT. He holds an M.Sc. in electrical engineering from KTH Royal Institute of Technology in Stockholm, Sweden.

Ari Keränen

◆ is an expert in IoT standards and protocols at Ericsson Research in Finland.. He joined the company in 2007 and has since worked with various internet technologies ranging from multimedia signaling and peer-to-peer systems to the IoT. He holds an M.Sc. in communications engineering from Aalto University in Helsinki, Finland.

Ben Smeets

◆ is a senior expert in trusted computing at Ericsson Research. He holds a Ph.D. in information theory from Lund University, Sweden, where he also serves as a professor. He joined Ericsson Mobile Communications in 1998, and started out working on security solutions for mobile phone platforms. Smeets is currently working on trusted computing technologies in connection with containers and secure enclaves.

John Fornehed

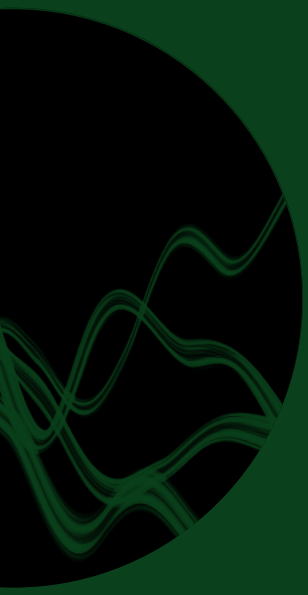
◆ joined Ericsson in 1991 and currently serves as an IoT expert and technical director. He spent many years in Japan, where he was responsible for strategic accounts with mobile operators, among other things. Fornehed's current work includes serving as an evangelist on IoT device life-cycle management, including secure IDs, for both industry and academia.

Carlos R. B. Azevedo

◆ joined Ericsson Research's Brazilian team in 2015. He currently serves as an MI and IoT technologies researcher at Ericsson Research in Stockholm, where he designs the architecture of intelligent, anticipatory and situation-aware systems. He holds a Ph.D. in electrical engineering from the University of Campinas in Brazil.

Peter von Wrycza

◆ joined Ericsson in 2011 and has held different positions in the areas of 3GPP standardization, 5G research and the IoT. He currently serves as head of IoT Technologies Research at Ericsson Research, where he drives the research, development and standardization activities for the IoT. Von Wrycza holds a Ph.D. in telecommunications from KTH Royal Institute of Technology in Stockholm.



ISSN 0014-0171
284 23-3324 | Uen

© Ericsson AB 2019
Ericsson
SE-164 83 Stockholm, Sweden
Phone: +46 10 719 0000