Critical IoT connectivity

IDEAL FOR TIME-CRITICAL COMMUNICATIONS

Critical Internet of Things (IoT) connectivity is ideal for a wide range of time-critical use cases across most industry verticals, and mobile network operators are uniquely positioned to deliver it.

**FREDRIK ALRIKSSON,
LISA BOSTRÖM,
JOACHIM SACHS,
Y.-P. ERIC WANG,
ALI ZAIDI**

Cellular Internet of Things (IoT) is driving transformation across various sectors by enabling innovative services for consumers and enterprises. There are currently more than one billion cellular IoT connections, and Ericsson forecasts that there will be around five billion connections by 2025 [1].

■ As 5G deployments gain momentum globally, enterprises in almost every industry are exploring the potential of 5G to transform their products, services and businesses. Since the requirements for wireless connectivity in different industries vary, it is useful to group them into four distinct IoT connectivity segments: Massive IoT, Broadband IoT, Critical IoT and Industrial Automation IoT [2].

While Massive IoT and Broadband IoT already exist in 4G networks, Critical IoT will be introduced with more advanced 5G networks. Industrial Automation IoT, the fourth segment, includes capabilities on top of Critical IoT that enable integration of the 5G system with real-time Ethernet and time-sensitive networking (TSN) used in wired industrial automation networks.

Critical IoT addresses the time-critical communication needs of individuals, enterprises and public institutions. It is intended for time-critical applications that demand data delivery within a specified time duration with required guarantee (reliability) levels, such as data delivery within 50ms with 99.9 percent likelihood (reliability).

Critical IoT is a paradigm shift from the enhanced mobile broadband (eMBB) connectivity, where the data rate is maximized without any guarantee on latency [3]. Many industry sectors have already started piloting time-critical use cases.

Time-critical use cases

The majority of time-critical use cases can be classified into the following four use case families:

- Industrial control
- Mobility automation
- Remote control
- Real-time media

Each family is relevant for multiple industries and includes a wide range of use cases with more or less stringent time-critical requirements, as shown in *Figure 1.*

Furthermore, there are three main network deployment scenarios depending on the coverage needs of time-critical services in different industries:

- Local area
- Confined wide area
- General wide area

Local-area deployment includes both indoor and outdoor coverage for a small geographical area such as a port, farm, factory, mine or hospital. Confined wide-area deployment is for a predefined geographical area – along a highway, between certain electrical substations, or within a city center, for example. General wide-area deployment is about serving devices virtually anywhere.

Common to all time-critical use cases is the fact that the communication service requirements depend on the dynamics of the use case and the application implementation. A highly dynamic system requires faster control with shorter round-trip times (RTTs), while a slower control loop is sufficient for a system that operates more slowly.

Various factors – such as device processing capabilities, the processing split between the device and the application server, the application's ability to extrapolate and predict data in case of missing

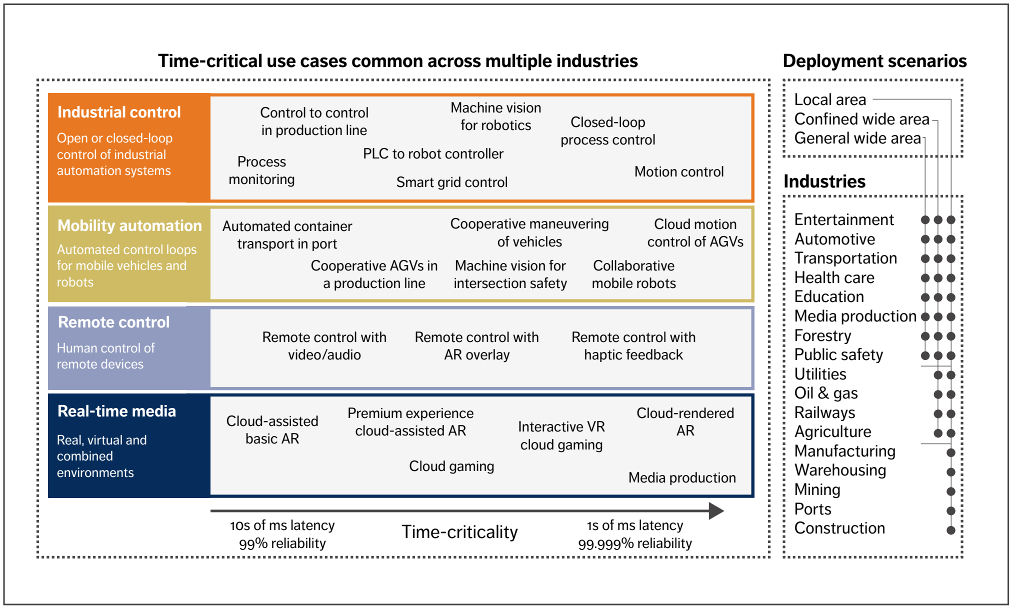

Figure 1 Examples of use cases enabled by Critical IoT

Time-critical use case trials

In partnership with leading industry partners and mobile network operators, Ericsson has trialed various Critical IoT use cases including:

- **Industrial control for manufacturing vehicles:** *https://www.ericsson.com/en/networks/cases/accelerate-factory-automation-with-5g-urllc*
- **Industrial control for manufacturing jet engines:** *https://www.youtube.com/watch?v=XZWC_ttighM*
- **Remote control in mining:** *https://www.youtube.com/watch?v=C4I0UKZ-FCc&t=7s*
- **Remote control of autonomous trucks:** *https://www.ericsson.com/en/press-releases/2018/11/ericsson-einride-and-telia-power-sustainable-self-driving-trucks-with-5g*
- **Remote bus driving:** *https://www.youtube.com/watch?v=IPyzGTD5FtM*
- **Cooperative vehicle maneuvers:** *https://5gcar.eu/*
- **Virtual reality and real-time media:** *https://www.ericsson.com/en/blog/2017/5/its-all-green-flags-for-5g-at-the-indianapolis-motor-speedway*
- **Augmented reality:** *https://www.ericsson.com/en/news/2018/3/5g-augmented-reality*
- **Smart harbor:** *https://www.ericsson.com/en/press-releases/2019/2/ericsson-and-china-unicom-announce-5g-smart-harbor-at-the-port-of-qingdao*

packets, rate adaptivity and which codecs are used – impact both the application RTT and the latency requirements on the communication network.

Industrial control includes a very broad set of applications, present in most industry verticals [4]. These applications typically consider late messages as lost. Process monitoring, controller-to-controller communication between production cells and some control functions for the electricity grid are examples of use cases with modest time-criticality, while use cases such as closed-loop process control and motion control have very stringent requirements.

Mobility automation refers to the automation of control loops for mobile vehicles and robots. Examples of the least time-critical use cases in this category include the relatively self-sufficient automated guided vehicles (AGVs) equipped with advanced on-board sensors that are used for transportation in ports and mines. Infrastructure-assisted vehicles such as fast-moving AGVs in a warehouse and collaborative maneuvering on public roads are examples of more time-critical mobility automation use cases, while the collaborative mobile robots used in flexible production cells represent an even higher degree of time-criticality.

Remote control refers to the remote control of equipment by humans. The ability to remotely control equipment is an important step in the evolution toward autonomous vehicles (to take temporary control of a driverless bus in scenarios not covered by its own automation functions) and for flying drones beyond visual line-of-sight.

Remote control can also improve work environments and productivity by moving humans out of inconvenient or hazardous environments – remote-controlled mining equipment [5] is one example. Such solutions also offer the benefit of providing enterprises with access to a broader workforce.

The communication service requirements for remote control depend on how fast the remote environment changes, the required precision of the task and the required QoE. Control-loop latency and audio/video quality are important factors for QoE and the ergonomics for the remote operator. Haptic feedback and augmented reality (AR) can be used to further improve the operator QoE and task precision, and will make the acceptable latencies even stricter.

Real-time media comprises use cases where media is produced and consumed in real time, and delays have a negative impact on QoE. Mobile applications for gaming and entertainment, including AR and virtual reality (VR), are common, with processing and rendering done locally in the device. Time-critical communication will make it possible to offload parts of the processing and rendering to the cloud [6], thereby improving the user experience and enabling the use of more lightweight devices (head-mounted, for example).

Time-critical communication can enable cloud gaming over cellular networks as well as new applications in sectors such as manufacturing, education, health care and public safety. It is expected to drive more widespread use of mobile AR and VR. Advanced media production (such as real-time production of live performances) with its strict delay and synchronization requirements, is another area where time-critical communication can enable new use cases.

Key network technologies and architectures

Achievable end-to-end (E2E) latencies depend on the available network and compute infrastructure, software features, and how the use case is implemented. In remote control, the physical distance between the remote operator and the teleoperated equipment is a physical property of the use case. In other use cases, the physical distance between end nodes can be reduced by distributed cloud processing, as in AR cloud gaming, where the AR overlay can be rendered in an edge cloud to limit interaction latencies. Network orchestration optimizes the placement of network and application functions to ensure efficient use of the compute and network infrastructure while restricting the transmission paths according to latency needs [7].

The 5G network comprises two functional domains: the next generation (5G) RAN (NG-RAN) and the 5G Core (5GC), which are built on an underlying transport network. All three – the NG-RAN, the 5GC and the transport network – contribute to the E2E reliability and latency, which is further affected by the device implementation.

The NG-RAN is deployed in a distributed fashion to provide radio coverage with good performance, availability and capacity. The 5GC provides connectivity of the device to the external services and applications. The network latency between the application and the RAN can be a major contributor to E2E latency.

Terms and abbreviations

5GC – 5G Core | **AAS** – Advanced Antenna System | **AGV** – Automated Guided Vehicle | **AR** – Augmented Reality | **CA** – Carrier Aggregation | **DC** – Data Center | **DL** – Downlink | **E2E** – End-to-End | **eMBB** – Enhanced Mobile Broadband | **FDD** – Frequency Division Duplex| **IoT** – Internet of Things | **MNO** – Mobile Network Operator | **NG-RAN** – Next Generation RAN | **NPN** – Non-Public Network | **NR** – New Radio | **PLC** – Programmable Logic Controller | **RTT** – Round-Trip Time | **TDD** – Time Division Duplex | **TSN** – Time-Sensitive Networking | **UE** – User Equipment | **UL** – Uplink | **URLLC** – Ultra-Reliable Low-Latency Communication | **VR** – Virtual Reality

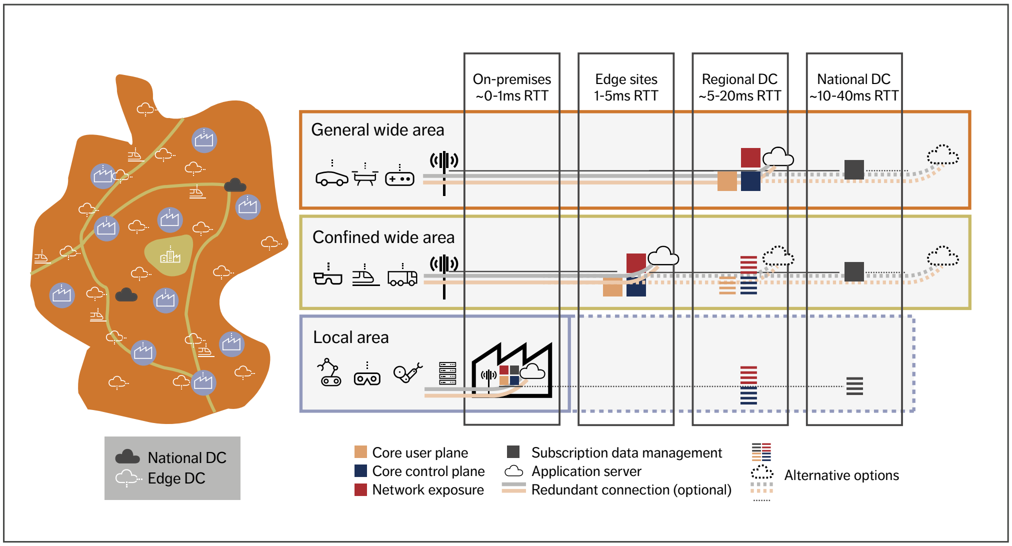

Figure 2 Examples of network architectures for low latency and/or high reliability

Figure 2 provides examples of network architectures for low latency and/or high reliability, and illustrates the effect of moving the application closer to the device. If an application is hosted in a central national data center (DC), the transport network round-trip latency can be in the order of 10-40ms, depending on the distance to the DC and how well the transport network is built out. Transport latency can be reduced to 5-20ms by moving applications to a regional DC or even to 1-5ms for edge sites. For local network deployments with networking functions and applications hosted on-premises, transport latencies become negligible.

Control of the network topology and the transport latency can be achieved by placing virtualized core network functions for execution at any location within the distributed computing platform of the network. This software-based design provides flexibility in updating the network with new functionality and reconfiguring it according to requirements. In addition to running telecommunication functions, the distributed computing platform allows the hosting of application functions in the network [8].

Network slicing makes it possible to create multiple logical networks that share a common network infrastructure. A dedicated network slice can be created by configuring and connecting computing and networking resources across the radio, transport and core networks. By reserving resources, a high availability of time-critical services can be ensured and latencies for queuing can be avoided.

Network orchestration automates the creation, modification and deletion of slices according to a slice service requirement [2]. This can imply that compute locations are selected according to guaranteed resource availability and transport latency rather than the lowest compute costs, for example. 5G New Radio (NR) provides several capabilities for ultra-reliable low-latency communication (URLLC) [7, 9]. From the first

NR standard release, the target has been to enable one-way latencies through the RAN of down to 1ms, where a timely data delivery can be ensured with 99.999 percent probability.

Features addressing low latency include ultra-short transmissions, instant transmission mechanisms to minimize the waiting time for uplink (UL) data, rapid retransmission protocols that minimize feedback delays from a receiver to the transmitter, instant preemption and prioritization mechanisms, interruption-free mobility and fast processing capabilities of devices and base stations.

Features addressing high reliability include a range of robust signal transmission formats. There are methods for duplicate transmissions to improve reliability through diversity, both within a carrier using transmissions through multiple antenna points, as well as between carriers through either carrier aggregation (CA) or multi-connectivity.

Advanced antenna systems (AAS) have tremendous potential to improve the link budget and reduce interference. The vendor-specific radio network configuration, algorithms for scheduling, link adaptation, admission and load control that are at the heart of NR make it possible to fulfill service requirements while ensuring an optimized utilization of available resources.

Support for highly reliable communication has also been addressed for the 5GC, by introducing options for redundant data transmission. Multiple redundant user-plane connections with disjoint routes and nodes can be established simultaneously. This may include the usage of separate user equipment (UE) on the different routes. 5G provides QoS, and by configuring a suitable QoS flow through the 5G system for transporting time-critical communication, queuing latencies due to conflicting traffic can be avoided by traffic separation with resource reservations and/or traffic prioritization.

For a time-critical communication service that is requested by a consumer, a data session with a suitable QoS flow profile is established, according to a corresponding service subscription. Larger customers, like an enterprise, are typically interested in connectivity for an entire device group. For this purpose, 5G has defined non-public networks (NPNs), which are real or virtual networks that are restricted for usage by an authorized group of devices for their private communication [10]. An NPN can be realized as a standalone network not coupled to a public network that is purpose-built to provide customer services at the customer premises.

Alternatively, an NPN may share parts of the network infrastructure with a public network, like a common RAN that is shared for private and public users. Beyond the shared RAN, the NPN may have a separate dedicated core network and local breakout – that is, it may be located on the customer's premises with its own device authentication, service handling and traffic management. Finally, an NPN can be a network service that is provided by a mobile network operator (MNO) as a customer-specific network slice.

Some NPNs may be customized to provide dedicated functionality for industrial automation, including 5G-LAN services and Ethernet support, providing ultra-low deterministic latency, interworking with IEEE (the Institute of Electrical and Electronics Engineers) TSN, and time-synchronization to synchronize devices over 5G to a reference time [11, 12]. Enhanced service exposure of the 5G system makes it possible to better integrate 5G into an industrial system [13] by means of service interfaces for device management (device onboarding, connectivity management and monitoring, for example) and network management.

> “… [AAS] HAVE TREMENDOUS POTENTIAL TO IMPROVE THE LINK BUDGET AND REDUCE INTERFERENCE ”

Spectrum option	Frequency allocation	Deployment scenario	Subcarrier spacing
Low-band FDD	2x10MHz @ 800MHz	Wide area	15kHz
Mid-band FDD	2x20MHz @ 2GHz	Wide area	15kHz
Mid-band TDD	50MHz @ 3.5GHz	Wide area	30kHz
Mid-band TDD	100MHz @ 3.5GHz	Local factory	30kHz
High-band TDD	400MHz @ 30GHz	Local factory	120kHz

Table 1 Spectrum assets considered in the case studies

5G spectrum flexibility

5G NR allows MNOs to take full advantage of all available spectrum assets. NR can be deployed using the spectrum assets used for the LTE networks, either through refarming or spectrum sharing [14]. Most of the LTE spectrum assets are in the low and mid bands, which in the 5G era will continue to be used for wide-area coverage. Traffic growth will drive the need for increased network capacity throughout the 5G era.

Increased capacity can be achieved by adding more spectrum assets, densifying the network and/or upgrading capabilities at existing sites. New 5G spectrum options in the mid bands (around 3.5GHz) and in the high bands (such as the millimeter wave frequencies) present great opportunities with large bandwidths.

Operating with these new spectrum assets, the added RAN nodes can also use advanced hardware features such as an AAS to fully capitalize on the benefits of NR. The coverage provided by the low-band and mid-band spectrum assets is key to enable Critical IoT services in wide-area deployments. Adding network capacity over time will not only increase the capacity for eMBB, but also boost the capacity for Critical IoT.

Case studies

To illustrate how 5G spectrum assets can be utilized for Critical IoT, we have put together case studies for two deployment scenarios: wide-area deployment and local-area deployment inside a factory.

The wide-area scenario is based on a macro-deployment in central London with an inter-site distance of approximately 450m, assuming low-band FDD, mid-band FDD and mid-band TDD spectrum options. For the mid-band deployments, we include an AAS, with eight antenna columns for 3.5GHz and four for 2GHz. Devices with four receiver branches are used in the evaluation.

The local factory setup is based on a factory automation scenario [15] and assumes mid-band and high-band TDD options. *Table 1* lists the spectrum options chosen in the case studies.

The top half of *Figure 3* presents the served capacity per cell versus various reliability and round-trip RAN latency requirements for outdoor UEs in the central London wide-area deployment scenario. All the TDD cases assume a TDD pattern with 3:1 downlink (DL) and UL split. Observe the cost in terms of capacity when pushing for tighter reliability and latency requirements. Generally, a tighter reliability or latency requirement leads to a higher consumption of radio resources, as the scheduler needs to provision a larger link adaptation margin to reduce the likelihood of failures in the initial transmissions. Furthermore, we observe that the mid-band options can offer a significant capacity boost for the wide-area scenario, thanks to large available bandwidths and use of AAS.

Among the two mid-band options studied, FDD at 2GHz is attractive when greater UL coverage (99 percent) is desired. Our case studies also show that it is challenging for the wide-area deployment to provide full indoor UL Critical IoT coverage using mid-band spectrum options, due to building-penetration loss. In general, indoor coverage depends on building materials and building sizes.

Under favorable conditions, such as low-loss facades and limited building sizes (that is, less than 3,600sq m in footprint), it is feasible to have 95 percent indoor UL coverage even using the mid-band carriers, although the achievable capacity is limited. Local indoor deployments are a prerequisite in high-loss or very large buildings, and are also necessary in other buildings if high indoor coverage and capacity is desired.

Although suburban and rural scenarios typically have larger cells, it is nonetheless possible to achieve similar results there. This is because antennas in suburban and rural environments tend to be installed at a greater height, there are fewer obstacles and the smaller buildings result in less wall-penetration loss. These factors compensate for the differences in cell range, making it feasible to achieve very good Critical IoT performance in suburban and rural scenarios as well.

For local-area studies (scenario #2), the deployment using 3.5GHz spectrum is based

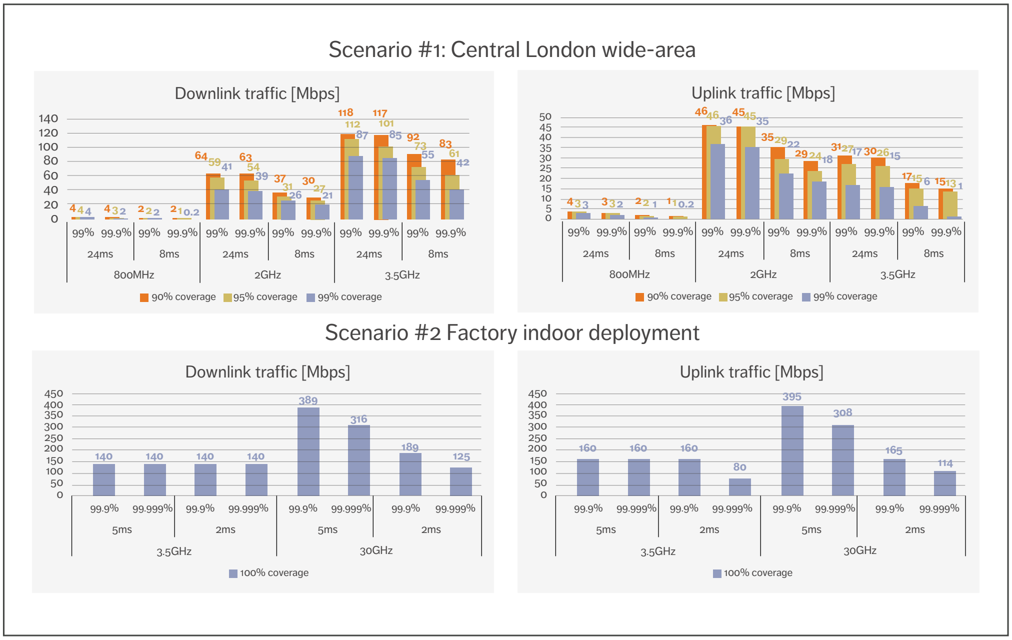

Figure 3 Served capacity per cell versus various reliability and round-trip RAN latency requirements for the two scenarios

> “ MNOs CAN START TO ADDRESS TIME-CRITICAL USE CASES ... THROUGH SOFTWARE UPGRADES ”

on a single-cell deployment with eight antennas installed in the ceiling, uniformly distributed across the entire factory, and a DL and UL symmetric TDD pattern. For the high-band deployment, eight transmission points with full frequency reuse are considered. The 3GPP indoor factory channel model with dense clusters, including machinery, assembly lines, storage shelves and so on [16], is used. To achieve 2ms round-trip RAN latency, NR mini-slot and configured grant features are used. (Using the same features, an FDD carrier with 15kHz subcarrier spacing can also achieve similar latency.)

The bottom half of Figure 3 shows that both DL and UL traffic achieve 100 percent coverage. Tightening Critical IoT requirements reduces capacity, however, and this is more evident in the high-band case. For the mid-band case, all users consistently reach the highest spectral efficiency except for UL traffic with the most stringent requirements, due to good coverage and the absence of interference achieved by the single-cell distributed antenna deployment.

5G NR CA allows radio resources from multiple carriers in multiple bands to be pooled to serve a user. For example, DL traffic can be delivered using a mid-band carrier even when the UL service requirements are not attainable on that mid-band carrier, by using a low-band carrier for the UL control and data traffic. This allows the DL capacity of the mid-band carrier to be utilized to a greater extent.

In essence, inter-band CA allows an MNO to improve coverage, spectral efficiency and capacity by dynamically directing the traffic through the better carrier, depending on the operating condition, user location and use case requirements. With a low-band carrier, there are also benefits of pooling an FDD carrier and a TDD carrier from a latency point of view, using the FDD carrier to mitigate the extra alignment delay introduced on a TDD carrier due to the DL-UL pattern.

Deployment strategy

MNOs have started to upgrade some 4G LTE radio base station equipment to 5G NR through software upgrades. The dynamic spectrum sharing solution allows efficient coexistence of LTE and NR in the same spectrum band down to millisecond level [14].

MNOs can start to address time-critical use cases in the wide area (the entertainment, health care and education sectors, for example) by adding support for Critical IoT connectivity to the NR carriers through software upgrades. More stringent, time-critical requirements call for radio network densification, edge computing, and further distribution and duplication of core network functions, which can be done gradually over time, while maximizing returns on investment.

In the confined wide-area scenarios (railways, utilities, public transport and the like), relatively stringent requirements can be addressed with reasonable investments in existing and new infrastructure. In local-area scenarios such as factories, ports and mines, even extreme time-critical requirements can be supported once the E2E ecosystem is established.

Dedicated spectrum has been allocated to some industry sectors in certain regions. In the wide-area scenarios such as public safety and railways, the allocated bandwidths are typically small (10MHz or below) and unable to meet the capacity demands of emerging use cases, especially those with time-critical requirements.

In some regions, significant TDD spectrum has been allocated to enterprises for local use (in the order of 100MHz) in mid-band and millimeter-wave frequency ranges. For both confined wide-area and local-area scenarios, the reuse of MNOs' existing infrastructure and their flexible spectrum assets (in combination with dedicated spectrum, if available) brings major value and opportunities. This approach makes it possible to exploit the full potential of various band combinations and support seamless mobility and interaction between public and dedicated communication infrastructure.

Conclusion

Critical Internet of Things connectivity addresses time-critical communication needs across various industries, enabling innovative services for consumers and enterprises. Mobile network operators are uniquely positioned to enable time-critical services with advanced 5G networks in a systematic and cost-effective way, taking full advantage of flexible spectrum assets, efficient reuse of existing footprint and flexible software-based network design.

Further reading

- **Ericsson, Evolving Cellular IoT for industry digitalization, available at:** *https://www.ericsson.com/en/networks/offerings/cellular-iot*
- **Ericsson, IoT connectivity, available at:** *https://www.ericsson.com/en/internet-of-things/iot-connectivity*

References

1. **Ericsson Mobility Report, November 2019, available at:** *https://www.ericsson.com/en/mobility-report/reports/november-2019*
2. **Ericsson white paper, Cellular IoT in the 5G era, February 2020, available at:** *https://www.ericsson.com/en/reports-and-papers/white-papers/cellular-iot-in-the-5g-era*
3. **3GPP TR38.913, Study on Scenarios and Requirements for Next Generation Access Technologies, 2017, available at:** *http://www.3gpp.org/ftp//Specs/archive/38_series/38.913/38913-e30.zip*
4. **5G-ACIA white paper, 5G for Automation in Industry – Primary use cases, functions and service requirements, July 2019, available at:** *https://www.5g-acia.org/fileadmin/5G-ACIA/Publikationen/5G-ACIA_White_Paper_5G_for_Automation_in_Industry/WP_5G_for_Automation_in_Industry_final.pdf*
5. **Ericsson Consumer and IndustryLab Insight Report, A case study on automation in mining, June 2018, available at:** *https://www.ericsson.com/en/reports-and-papers/consumerlab/reports/a-case-study-on-automation-in-mining*
6. **GSMA, Cloud AR/VR Whitepaper, May 8, 2019, available at:** *https://www.gsma.com/futurenetworks/resources/gsma-online-document-cloud-ar-vr-whitepaper/*
7. **Proceedings of the IEEE, vol. 107, Issue 2, pp. 325-349, Adaptive 5G Low-Latency Communication for Tactile Internet Services, February 2019, Sachs, J. et al., available at:** *http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8454733&isnumber=8626773*
8. **Ericsson Technology Review, Distributed cloud – a key enabler of automotive and industry 4.0 use cases, November 20, 2018, Boberg, C; Svensson, M; Kovács, B, available at:** *https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/distributed-cloud*
9. **Academic Press, Cellular Internet of Things – From Massive Deployments to Critical 5G Applications, October 2019, Liberg, O; Sundberg, M; Wang, E; Bergman, J; Sachs, J; Wikström, G, available at:** *https://www.elsevier.com/books/cellular-internet-of-things/liberg/978-0-08-102902-2*
10. **NGMN white paper, 5G E2E Technology to Support Verticals' URLLC Requirements, November 18, 2019, available at:** *https://www.ngmn.org/publications/5g-e2e-technology-to-support-verticals-urllc-requirements.html*
11. **Ericsson Technology Review, Boosting smart manufacturing with 5G wireless connectivity, January 2019, available at:** *https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/boosting-smart-manufacturing-with-5g-wireless-connectivity*
12. **Ericsson Technology Review, 5G-TSN integration meets networking requirements for industrial automation, August 2019, Farkas, J; Varga, B; Miklós, G; Sachs; J, available at:** *https://www.ericsson.com/en/ericsson-technology-review/archive/2019/5g-tsn-integration-for-industrial-automation*
13. **5G-ACIA, Exposure of 5G Capabilities for Connected Industries and Automation Applications (white paper), June 2020, available at:** *https://www.5g-acia.org/publications/*
14. **Ericsson Spectrum Sharing, available at:** *https://www.ericsson.com/en/networks/offerings/5g/sharing-spectrum-with-ericsson-spectrum-sharing*
15. **3GPP TR 38.824, Study on physical layer enhancements for NR URLLC, 2019, available at:** *http://www.3gpp.org/ftp//Specs/archive/38_series/38.824/38824-g00.zip*
16. **3GPP TR 38.901, Study on channel model for frequencies from 0.5 to 100 GHz, 2019, available at:** *http://www.3gpp.org/ftp//Specs/archive/38_series/38.901/38901-g00.zip*

THE AUTHORS

Fredrik Alriksson

◆ is a researcher at Development Unit Networks, where he leads strategic technology and concept development within IoT & New Industries. He joined Ericsson in 1999 and has worked in R&D with architecture evolution covering a broad set of technology areas including RAN, Core, IMS and VoLTE. Alriksson holds an M.Sc. in electrical engineering from KTH Royal Institute of Technology in Stockholm, Sweden

Lisa Boström

◆ is a researcher at Development Unit Networks, where she does research and concept development within IoT & New Industries. She joined Ericsson in 2006 and has worked extensively with RAN R&D and standardization. Boström holds an M. Sc. in media engineering from Luleå University of Technology in Sweden.

Joachim Sachs

◆ is a principal researcher at Ericsson Research in Stockholm and coordinates research activities on 5G for industrial IoT solutions and cross-industry research collaborations. He holds a Ph.D. from the Technical University of Berlin in Germany. Sachs is coauthor of the book Cellular Internet of Things: From Massive Deployments to Critical 5G Applications.

Y.-P. Eric Wang

◆ joined Ericsson in 1995 and is currently a principal researcher at Ericsson Research. He holds a Ph.D. in electrical engineering from the University of Michigan (Ann Arbor) in the US. Wang is coauthor of the book Cellular Internet of Things: From Massive Deployments to Critical 5G Applications.

Ali Zaidi

◆ is a strategic product manager for Cellular IoT at Ericsson and also serves as the company's head of IoT Competence. He holds a Ph.D. in telecommunications from KTH Royal Institute of Technology in Stockholm. Since joining Ericsson in 2014, he has been working with technology and business development of 4G and 5G radio access at Ericsson. Zaidi is currently responsible for LTE-M, URLLC, Industrial IoT, vehicle-to-everything and local industrial networks.

The authors would like to thank Yanpeng Yang, Anders Furuskär, Kittipong Kittichokechai, Anders Bränneby, Fedor Chernogorov, Gustav Wikström, Jari Vikberg, Mattias Andersson, Ralf Keller, Kun Wang, Torsten Dudda and Marie Hogan for their contributions to this article.