5G AMERICAS WHITE PAPER

5G – THE FUTURE OF IOT

JULY 2019

CONTENTS

EXECUTIVE SUMMARY

5G: The Future of IoT takes a look at the market drivers, trends and cellular technology solutions that will create our connected future.

Market drivers provide added value through connectivity of all "things" ranging from street lighting to home appliances to industrial robotics. Providing connectivity to things has become easier with improvements in the economics of end devices, large investments in IoT systems, adoption of global standards and availability of spectrum. Overall trends in information technology such as cloud computing and edge cloud, artificial intelligence and security assurance have accelerated the IoT ecosystem. The whitepaper discusses some of the market segments in more details, including industrial IoT, smart cities, enterprise IoT and consumer IoT.

The whitepaper also provides an overview of developments in 3GPP standards for IoT, looking at both the Massive IoT and Critical IoT segments. It starts with the 3GPP requirements for IoT and goes on to cover the high level 3GPP solutions to meet these requirements. For Critical IoT, the Ultra-Reliable Low Latency Communications (URLLC) radio and architecture are described. For Mission Critical IoT, it describes the role of the 5G core network as well as improvements in the radio. It also addresses the factory automation vertical for which several new features have been introduced in 3GPP to allow operation of 5G systems together with IEEE TSN. It provides an update of spectrum options for licensed, unlicensed and shared bands. In essence, the 3GPP standards are themselves a market driver in the future-proofing and delivery of IoT as networks transform from 4G to 5G.

The 5G networks being deployed today will build upon the 4G LTE networks including both Narrowband-IoT (NB-IoT) and LTE Category M1 (LTE-M or LTE for Machines). Additional security, automation and management functions on 5G radio and core networks will deliver the ultra-reliability requirements for Critical IoT solutions. As use cases continue to develop in the market with more applications, the Industrial IoT, Consumer IoT and Enterprise IoT markets are flourishing. The massive IoT of the future –with connectivity across many vertical domains—will be insured by the mobile cellular Internet of Things.

1. INTRODUCTION

The Internet of Things (IoT) is transforming businesses and peoples' lives, and will continue to ignite innovations in the future. It is expected that we will progress to tens of billions of connected devices globally over the next decade[1] that will generate multi-trillion dollars of economic value[2] across many markets forming the foundation of a totally interconnected world—or the Internet of Everything. In this vision of a totally connected world, cellular technologies will play a pivotal role – and they already have; 1G and 2G networks connected people to one another via voice, and 3G and 4G extended connectivity to the mobile Internet with fast mobile broadband services.

[1] *Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016*, Gartner. 7 February 2017.
[2] *IDC Forecasts Worldwide Spending on the Internet of Things to Reach $772 Billion in 2018*, IDC. 7 December 2017.

There are already more connected things than people in the world. Analyst firm Gartner reported that about 8.4 billion IoT devices were in use in 2017, up 31 percent from 2016, and this will likely reach 20.4 billion by 2020.[3]

Another leading analyst firm, IDC, put worldwide spending on IoT at $772.5 billion in 2018 -- up nearly 15 percent on the $674 billion spent in 2017. IDC predicts that total spending will hit $1 trillion in 2020 and $1.1 trillion in 2021.[4]

Today's IoT devices are connected on a wide variety of wireless technologies that might roughly fall into four classes of connectivity technologies: wired, short to mid-range wireless (from Bluetooth to mesh networking), long-range wireless (including cellular and Low Power Wide Area networking (LPWA)), and satellite. Within each class are numerous specific technologies and standards. Another way to consider the IoT technology classes is by dividing them into short-range and wide-area connectivity segments. The former is typically enabled by unlicensed radio technologies, such as Wi-Fi, Bluetooth, ZigBee and Z-wave. The latter consists of devices powered by cellular technologies (GSM, LTE, and 5G) as well as unlicensed low-power technologies, such as Sigfox, and LoRA. These technologies offer different benefits and use cases. For the purpose of this report, the focus remains on the wireless cellular technologies of the global Third Generation Partnership Project (3GPP) standards, and in particular the foundation of 4G LTE for the future of 5G and the massive IoT.

To be clear, "massive IoT" refers to the tens of billions of devices, objects and machines that require ubiquitous connectivity, whether mobile, nomadic or stationary. To reach the massive scale as defined by 3GPP that would mean at least one million devices per kilometer (km). In addition, the mobile networks must efficiently support the simplest devices communicating infrequently that are ultra-energy efficient for the long term with ten-plus years' battery life. However, the massive IoT does not have to wait for 5G scale deployments for success. 4G technology is supporting the massive IoT today.

According to the Ericsson Mobility Report - Q4 2018, there will be approximately 400 million IoT devices with cellular connections at the end of 2016 and that number is projected to reach 1.5 billion in 2022, or around 70 percent of the wide-area category. In 2018, mobile phones connections were surpassed in numbers by IoT devices.[5] This growth is due to increased industry focus and 3GPP standardization of cellular IoT technologies.

Cellular IoT connections benefit from enhancements in provisioning, device management and service enablement. Perhaps even more significantly, cellular networks also offer ubiquitous coverage, and unparalleled levels of reliability, security and performance required by the most demanding IoT applications. 3GPP technologies such as 4G Long Term Evolution (LTE) can provide wide-area IoT connectivity—LTE is the most widely deployed and fastest growing wireless technology with 3.7 billion connections worldwide at the end of 2018.[6] With 637 LTE commercial networks[7] worldwide as of April 2019, LTE will continue to expand its' coverage and connections over the next decade forming the foundation for the next generation of wireless technology – 5G.

3GPP has already introduced a suite of two complementary narrowband LTE IoT technologies in Release 13: eMTC (enhanced Machine-Type Communication), also known as LTE-M (Machine-Type Communication) and NB-IoT (NarrowBand-Internet of Things) often collectively referred to as LTE IoT. Both are optimized for lower complexity/power, deeper coverage, and higher device density, while seamlessly

[3] *Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016*, Gartner. 7 February 2017.
[4] *IDC Forecasts Worldwide Spending on the Internet of Things to Reach $772 Billion in 2018*, IDC. 7 December 2017.
[5] *Ericsson Mobility Report*, *Internet of Things Forecast*, Q4 2018.
[6] WCIS, Ovum. March 2019.
[7] TeleGeography. April 2019.

coexisting with other LTE services such as regular mobile broadband. LTE IoT starts to connect the massive IoT today.

Figure 1.1. Complementary eMTC & NB-IoT Address a Wide Range of IoT Use Cases.[8]

LTE-M is the industry term for the Long-Term Evolution (LTE) machine-type communications (MTC) LPWA technology standard. LTE-M supports lower device complexity, massive connection density, low device power consumption, low latency and provides extended coverage, while allowing the reuse of the LTE installed base. The deployment of LTE-M can be done "in-band" within a normal LTE carrier, or "standalone" in a dedicated spectrum.

Narrowband IoT (NB-IoT) is a 3GPP radio technology standard that addresses the LPWA requirements of the IoT. NB-IoT is characterized by improved indoor coverage, support of massive number of low throughput devices, low delay sensitivity, ultra-low device cost, low device power consumption and optimized network architecture. Like LTE-M, NB-IoT can be deployed "in-band" within a normal LTE carrier, or "standalone" for deployments in dedicated spectrum. Additionally, NB-IoT can also be deployed in an LTE carrier's guard-band.

[8] *LTE IoT is starting to connect the massive IoT today, thanks to eMTC and NB-IoT,* OnQ Blog, Hao Xu, Qualcomm Principle Engineer.15 June 2017.

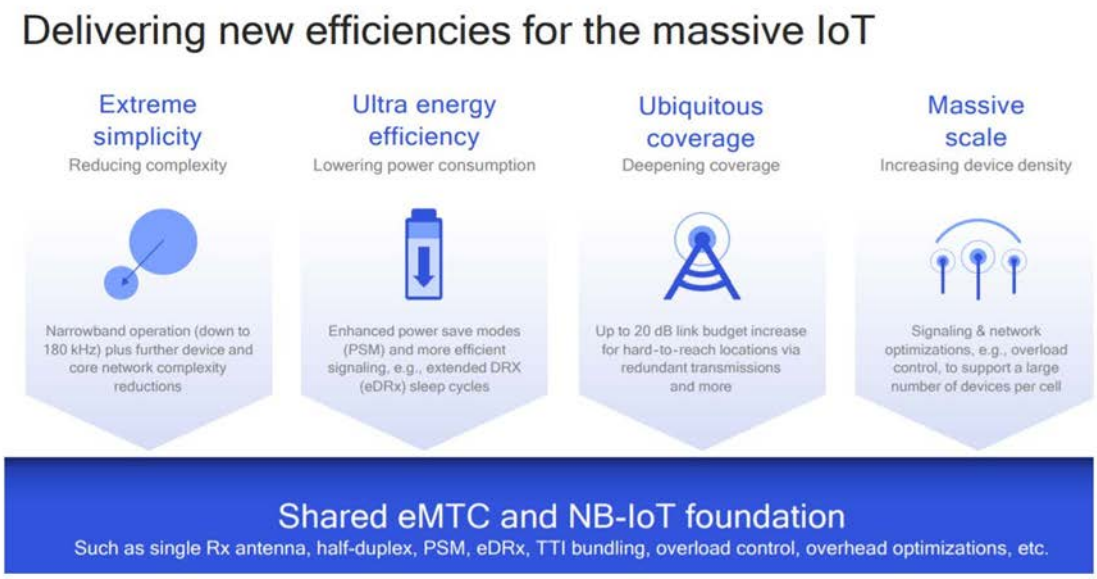


Figure 1.2. Delivering New Efficiencies for the Massive IoT.[9]

Beyond 3GPP Release 13, there is a future roadmap of LTE IoT technology innovations that deliver additional enhancements to meet tomorrow's massive IoT requirements.

For example:

- Release 14 brings single-cell multicast for easy over-the-air firmware upgrades and device positioning for asset location tracking
- Release 15 introduces Time Division Duplex (TDD) support for NB-IoT, as well as a new wake-up receiver design to allow for even better energy efficiency
- Release 16 offers enhancements for both LTE IoT and 5G NR IoT such as Non-Orthogonal Multiple Access (NOMA) enabled by Resource Spread Multiple Access (RSMA) that can further increase device density and network efficiency. Grant-free uplink will allow IoT devices to send sporadic small data bursts to the network without scheduling, thereby reducing overhead for more efficient handling of IoT communication. Another area in development in Rel-16 is mesh networking with Wide Area Network (WAN) management, which helps with extending range and optimizing device cost.

[9] *Leading the LTE IoT evolution to connect the massive Internet of Things*, Qualcomm. July 2018.

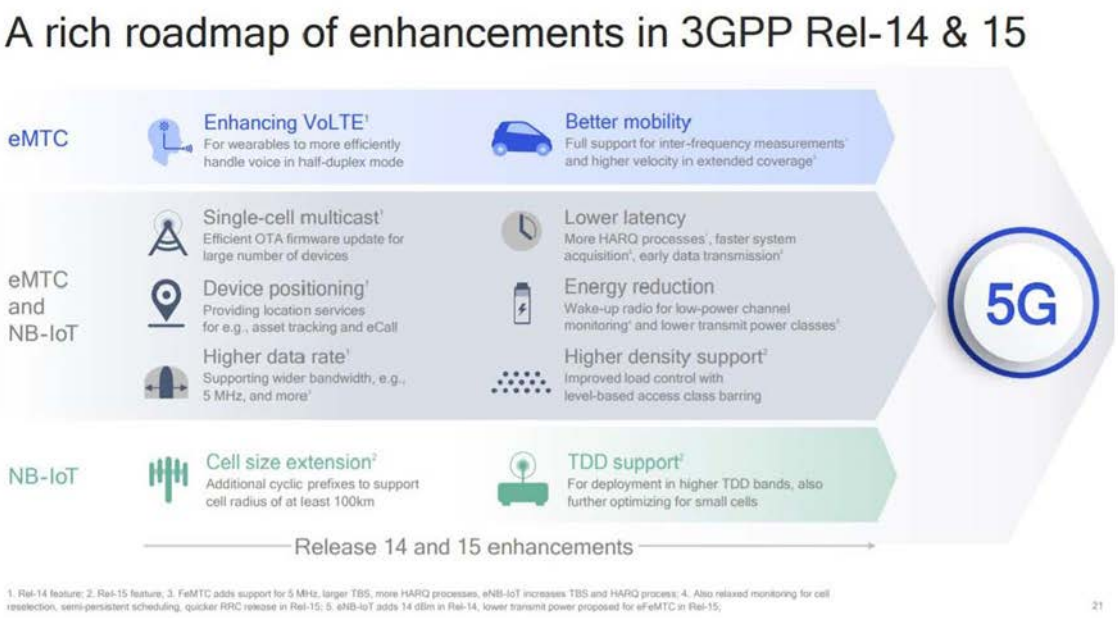


Figure 1.3. A Rich Roadmap of Enhancements in 3GPP Rel-14 & 15.[10]

As of July 2019, 68 mobile operators[11] have NB-IoT and/or LTE-M deployed on their networks to connect the massive IoT.

In 2019, as 5G NR eMBB services launch in both millimeter Wave (mmWave) and sub 6 Gigahertz (GHz) spectrum, LTE IoT will continue to evolve and operate seamlessly with the new 5G network. For example, LTE IoT is agnostic to core networks — either LTE core (EPC) or 5G core (NextGen or 5GC), will support LTE-IoT evolution. And LTE IoT deployments will be supported by 5G NR in new 5G spectrum bands, allowing 5G NR IoT to fully leverage LTE IoT investments. More information is provided later in this paper regarding the use of spectrum assets for IoT applications.

LTE IoT will continue to evolve over coming years, leveraging the scale, longevity and global coverage of LTE networks to not only seamlessly enable migration from 2G, but to also complement the initial 5G NR (New Radio) deployments that focus on enhanced mobile broadband and high-performance IoT. This continued evolution and its expanded deployments are integral parts of the 5G platform – a unified, more capable connectivity fabric for our future.

This white paper builds upon earlier publications by 5G Americas, including LTE and 5G Technologies Enabling the Internet of Things published in December 2016[12] and LTE Progress Leading to the 5G Massive Internet of Things published in December 2017.[13]

[10] *Leading the LTE IoT evolution to connect the massive Internet of Things*, Qualcomm. July 2018.
[11] *5G Americas – Cellular IoT Deployments.* Source: TeleGeography. July 2019.
[12] *LTE and 5G Technologies Enabling the Internet of Things*, 5G Americas white paper. December 2016.
[13] *LTE Progress Leading to the 5G Massive Internet of Things*, 5G Americas white paper. December 2017.

2. IOT MARKET OVERVIEW

IoT, initially referred to as Machine-to-Machine (M2M), represents one of the key growth opportunities for telecommunication service providers and enterprises of various sizes in the next decade. IoT represents all devices that are connected to the internet and can communicate with other connected devices through wireless networks and embedded sensors, so it has an expanded market from M2M that has evolved over time and with technology development. That evolution will continue; whereas 4G has been driven by device proliferation and dynamic information access, 5G will be driven largely by IoT applications, a wide range of IoT use cases, Massive IoT (MIOT) deployment as well as more advanced solutions that may be categorized as Critical IoT. Figure 2.1 displays the applications for both M-IoT and Critical IoT technology and the differing requirements that are satisfied by each. Further information on these applications and technologies are offered throughout the whitepaper.

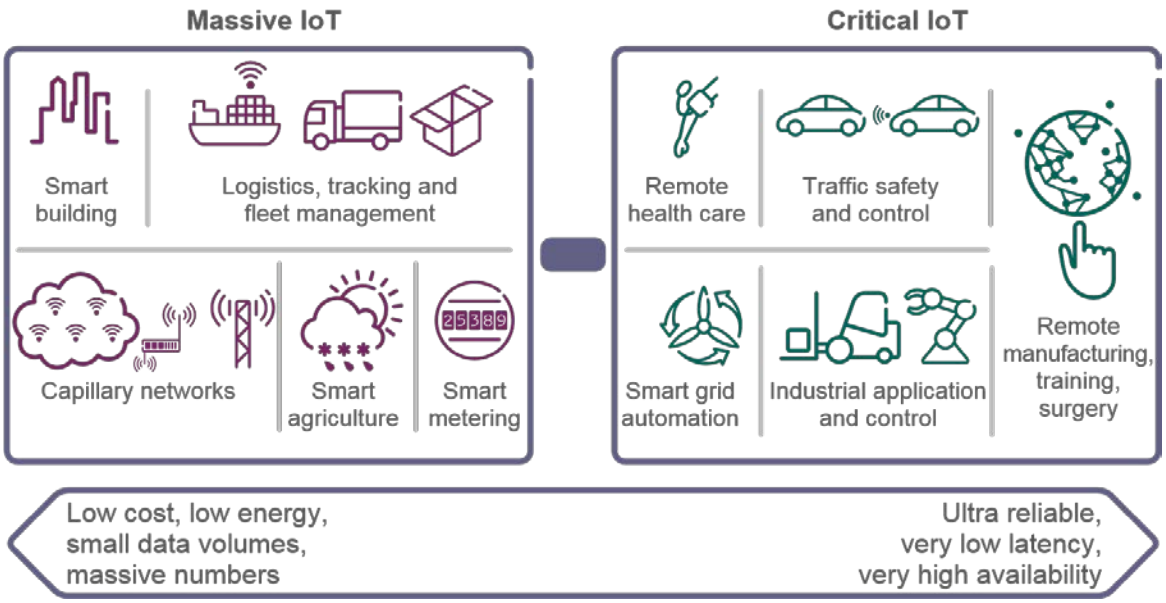


Figure 2.1. Differing Requirements for Massive and Critical IoT Applications.[14]

As the name Massive IoT suggests, 5G technology will enable a connected world numbering many billions of IoT devices. For example, research from Business Insider projects more than 64 billion IoT devices by 2025, up from about 10 billion in 2018, and 9 billion in 2017.[15] The massive IoT is further demonstrated by Cisco predicting that by 2022 there will be 14.6 billion machine-to-machine IoT connections as demonstrated in Figure 2.2.

[14] *Cellular Networks for Massive IOT*, Ericsson white paper. January 2016.
[15] *IoT Report: How Internet of Things Technology Growth is Reaching Mainstream Companies and Consumers*, Business Insider, Peter Newman. 28 January 2019.

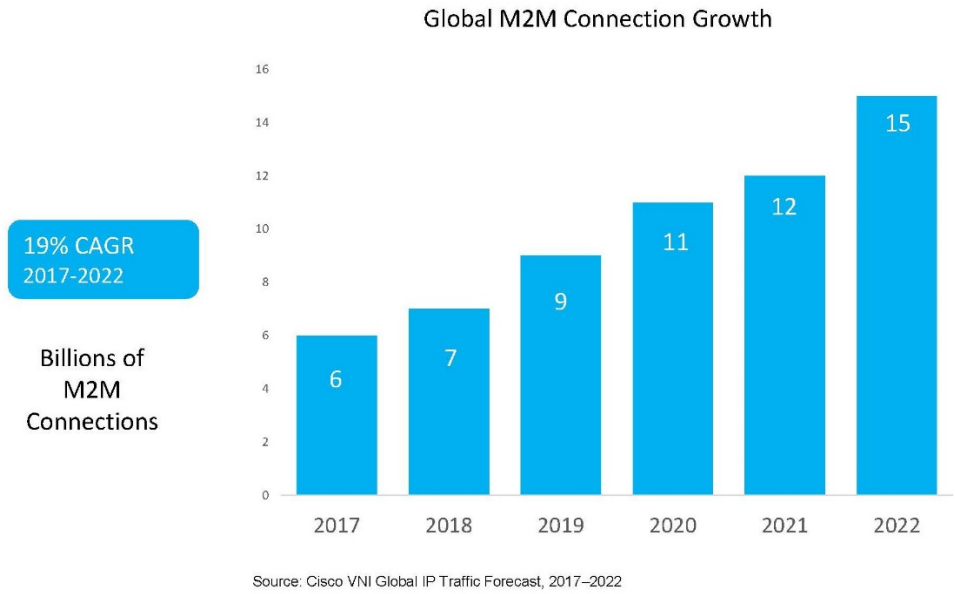


Figure 2.2. Global M2M Connection Growth.

2.1 5G AND IOT MARKET GROWTH

With 5G deployments starting in 2019, the mobile ecosystem is becoming even larger, and more widespread and extensive. Momentum is building in many markets as service providers accelerate their plans for 5G rollout. In 2024, Ericsson projects that 5G will reach 40 percent population coverage and 1.5 billion subscriptions, making it the fastest generation ever to be rolled out on a global scale. This is driven by new, innovative solutions that reuse existing infrastructure and available spectrum. In parallel to the 5G rollout, cellular IoT is passing new milestones on its way to becoming the technology of choice for wide-area IoT applications. Boosted by a strong uptake in North East Asia, Ericsson predicts that cellular IoT connections are set to pass the 4 billion mark by 2024.[16] As explained in the Introduction, the IoT will comprise numerous technologies, including wireless technologies. This paper will focus largely on wireless Cellular IoT (C-IoT).

5G roll-outs provide mobility innovation and new levels of fixed/mobile convergence that will be impactful in driving the IoT market. Cisco forecasts that by 2022, 22 percent of global Internet traffic will come from mobile (cellular) networks (up from 12 percent in 2017); also by 2022, about 3 percent of global mobile devices/connections will be 5G-capable and nearly 12 percent of global mobile traffic will come from 5G.[17] Mobile carriers from around the world are beginning to introduce trial and commercial 5G networks with large-scale 5G deployments beginning in 2020, when mobile spectrum, standards, profitable business plans and other operational issues are more fully developed. As of May 2019, there were commercial launches of 8 networks worldwide offering 3GPP standardized 5G technology, according to research firm TeleGeography.[18] That number is anticipated to grow to about 25 commercial 5G networks worldwide by the end of 2019.[19]

[16] *Ericsson Mobility Report*, letter from the publisher Fredrik Jejdling. November 2018.
[17] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.
[18] 5G Americas maintains a list of 5G deployments researched by TeleGeography at 5gamericas.org.
[19] *The 5G new network arrives,* Deloitte Insights. December 2018.

5G mobile devices and connections will be over 3 percent of 12.3 billion global mobile devices and connections by 2022 – meaning over 422 million devices will be 5G capable. Nearly twelve percent of global mobile traffic will be on 5G cellular connectivity by 2022 generating 21 GB of traffic per month.[20]

2.2 MARKET DRIVERS

IoT is the new game changer for businesses and individuals, and is considered to be the fourth industrial revolution by experts and scientists.[21] However, a revolution is often triggered by the rise and demand of certain driving factors. Similarly, IoT and its development are also backed by certain key components. Thus, if organizations want to prosper with IoT, they need to consider who and what is driving IoT and its innovation.

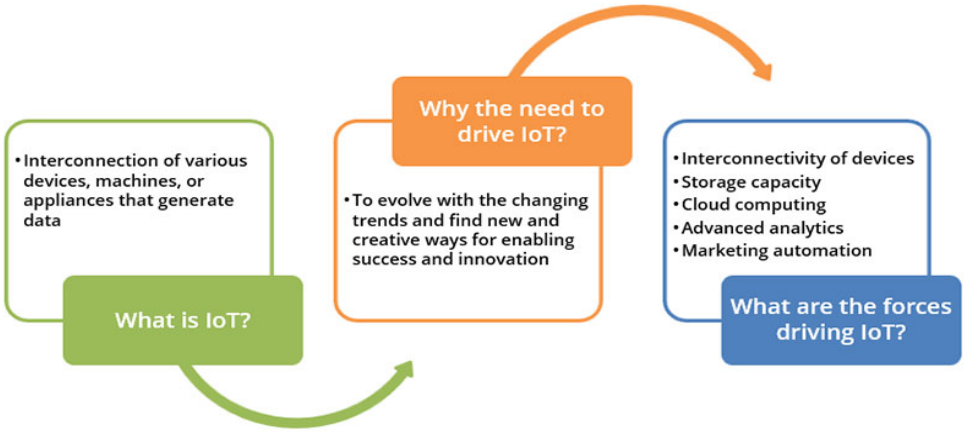


Figure 2.3. IoT Basically Defined.[22]

Market drivers are the underlying forces and trends that make markets develop and grow. What are the drivers for the IoT? What are the trends that are impacting the mobile industry and also changing and expanding the IoT?

The drivers and trends that are contributing to the development of the IoT market defined in this whitepaper include: 3GPP standards; expanded internet connectivity; high mobile adoption; low-cost sensors; large IoT investments; global application trends; emergence of new mobile technologies; growing importance of automation and big data knowledge; Artificial Intelligence (AI) and Machine Learning (ML); Edge Computing and the Cloud; more use cases across vertical domains; security assurance; IPv6; and Open Source for 5G (see Figure 2.4).

[20] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.
[21] Klaus Schwab, founder and executive chairman of the Geneva-based World Economic Forum, published a book in 2016 titled *The Fourth Industrial Revolution* and coined the term in Davos that same year.
[22] *What's Driving IoT?* Blog by Naveen Joshi, Allerin. 17 November 2017. https://www.allerin.com/blog/whats-driving-iot.

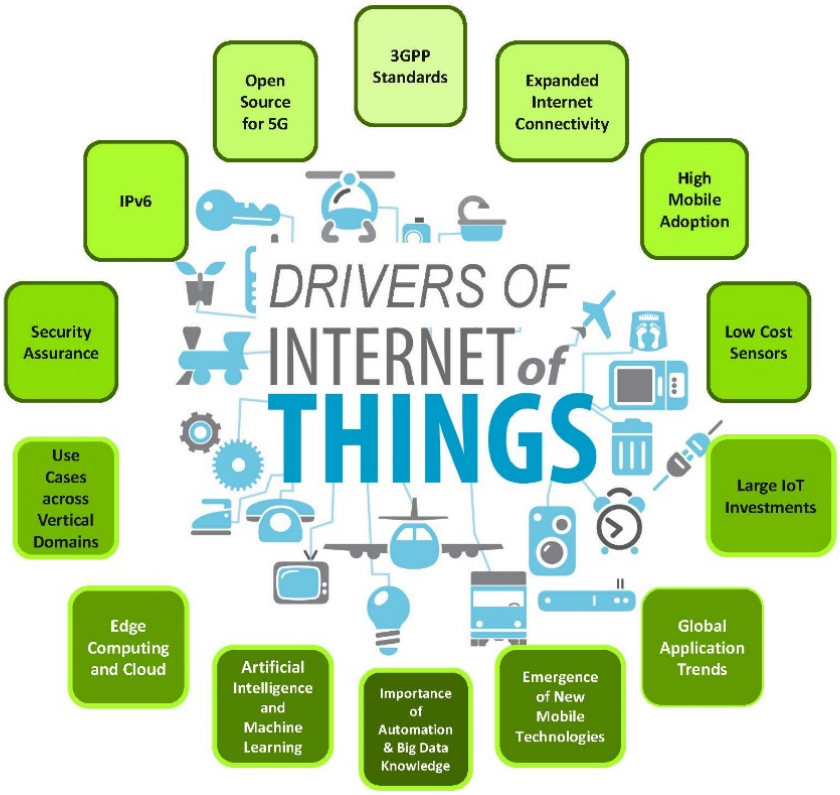


Figure 2.4. IoT Market Drivers and Trends.

2.2.1 EXPANDED INTERNET CONNECTIVITY

More and more devices—including cellular M2M devices—are being connected to the Internet. The IoT is no longer a phenomenon; it has become a prevalent system. As show in Figure 2.2 from Cisco's VNI, globally M2M connections will grow 2.4-fold, from 6.1 billion in 2017 to 14.6 billion by 2022. There will be 1.8 M2M connections for each member of the global population by 2022.

According to Cisco's forecast, the total number of devices connected to IP networks will be more than three times the global population by 2022. In the Cisco VNI the following forecasts are established by 2022:[23]

- 3.6 networked devices per capita, up from 2.4 networked devices per capita in 2017
- 28.5 billion networked devices by 2022, up from 18 billion in 2017
- M2M connections will be more than half of the global connected devices and connections
- Share of M2M connections will grow from 34 percent in 2017 to 51 percent
- 14.6 billion M2M connections by 2022

Among the countries that will have the highest average of per capita devices and connections by 2022 are the United States (13.6), South Korea (11.8), and Canada (11.0) compared to the global average of 3.6 networked devices per capita.

[23] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.

Each year, various new devices in different form factors with increased capabilities and intelligence are introduced and adopted in the market. A growing number of M2M applications, such as smart meters, video surveillance, healthcare monitoring, transportation, and package or asset tracking, are contributing in a major way to the growth of devices and connections. By 2022, M2M connections will be the fastest-growing category reaching 51 percent share of the total devices and connections.[24]

The number of cellular connected devices has grown at a compounded annual rate of 33 percent since 2013,[25] but growth is not only limited to the number of devices. Over the same period, traffic per connected device has grown much faster due to an increasing share of devices generating higher traffic volumes.

2.2.1.1 IOT TRAFFIC CHARACTERISTICS

Although traffic from all connected devices has grown significantly, the percentage of cellular IoT traffic still represents a very small portion of total mobile traffic in service providers' networks. Most of today's cellular IoT applications generate relatively small data traffic volumes in mobile networks. The reason is that the installed base of IoT devices is a distribution of 2G, 3G and LTE technology; a majority of these are 2G devices, due to the long-life cycles of sensors and applications with basic requirements. However, in the future, this distribution is expected to change as a broader range of use cases evolve over time, along with the continued deployment of supporting LTE-based IoT technologies and future capabilities of 5G networks.

2.2.1.2 IOT TRAFFIC VOLUME

As noted, IoT traffic volume is limited but increasing. So far, IoT has been characterized by a very large number of connections, small data volumes and, in some cases, stringent requirements on energy consumption. Typical uses are sensor, monitor or control data IoT applications and typically these were the first services to be built on NB-IoT technology. Data traffic generated by such devices is generally low; the typical data packet for a sensor-based service is about 100–150 bytes, with a payload comprised of a device ID, time stamp and reported data values. NB-IoT technology is capable of supporting data rates of 227 Kbps in uplink and 250 Kbps in downlink.[26] Cat-M1 or LTE-M is a second technology designed and standardized for massive IoT applications and is capable of supporting data rates up to 1 Mbps in both the uplink and downlink.

The traffic volume generated by massive IoT applications is a function of message size, message interval and the number of devices deployed per square kilometer. Beyond massive deployments of devices generating limited data volumes, there is an evolving range of IoT applications that have stringent requirements on availability, delay and reliability. Applications include traffic safety, automated vehicles, drones and industrial automation. These can generate many times more data traffic than massive IoT applications, depending on the specific use case. They could be based on LTE devices where, for example, an LTE Cat 4 device supports data rates of 150 Mbps in downlink and 50 Mbps in uplink; or have use case requirements that only future 5G network capabilities can meet. NB-IoT and Cat-M1 will continue to coexist with the introduction of 5G networks.

[24] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.
[25] *Ibid.*
[26] *Ericsson Mobility Report.* November 2018.

Traffic from WiFi and mobile devices will account for 71 percent of total IP traffic by 2022, according to Cisco. Globally, mobile data traffic will increase sevenfold between 2017 and 2022 and will grow from 9 percent to 20 percent of total IP traffic from 2017 to 2022.[27]

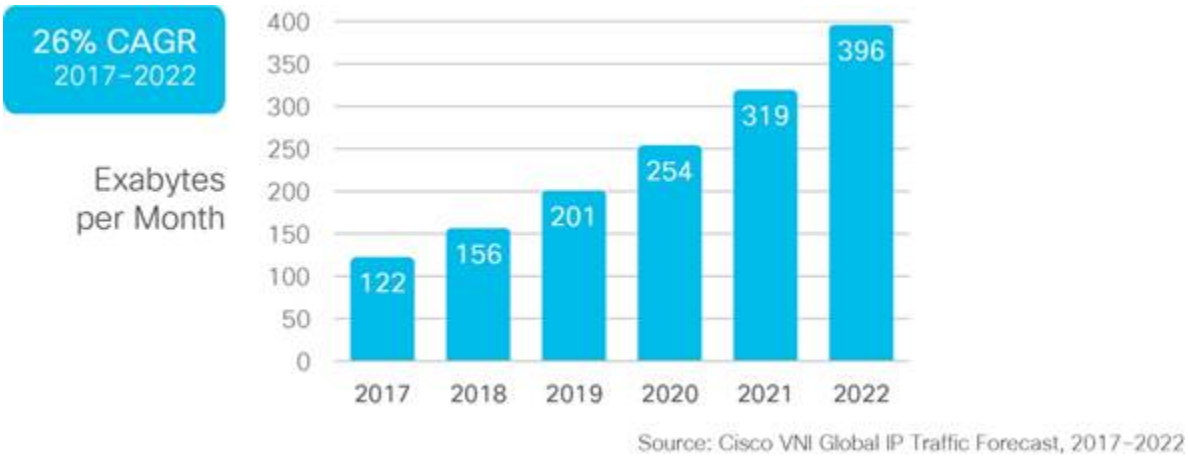


Figure 2.5. Monthly Global IP Traffic 2017-2022.[28]

Although the number of connections is growing 2.4-fold, global M2M IP traffic will grow more than sevenfold over this same period, from 3.7 EB per month in 2017 (3 percent of global IP traffic) to more than 25 EB by 2022 (6 percent of global IP traffic).[29] The amount of traffic is growing faster than the number of connections because of the increase of deployment of video applications on M2M connections and the increased use of applications, such as telemedicine and smart car navigation systems, which require greater bandwidth and lower latency.

2.2.2 HIGH MOBILE ADOPTION

The number of devices connected to IP networks will be more than three times the global population by 2022, according to Ericsson.[30] Cellular IoT connections will grow from one billion in 2018, growing to more than two billion by 2021 and more than four billion by 2024, as shown in Figure 2.6.

[27] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.
[28] *Ibid.*
[29] *Ibid.*
[30] *Ericsson Mobility Report.* November 2018.

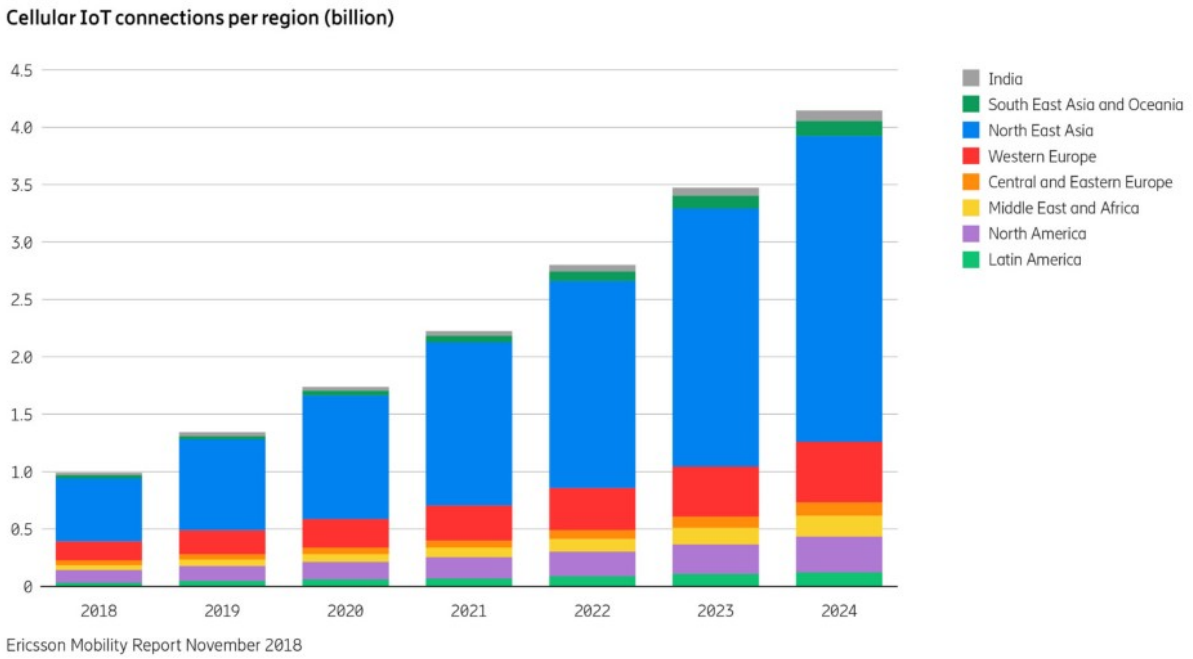


Figure 2.6. Cellular IoT Connections 2019-2024 by World Region.[31]

2.2.3 LOW COST SENSORS

Sensors are becoming prolific; the larger the scale, the lower the cost. The cost of sensors, actuators, transducer systems and declining hardware costs overall are resulting in a lower cost of entry for vertical markets, enterprises, consumer devices and others to the IoT.

And it is not only the lower costs of sensors and similar 'bits and pieces' of the IoT device that has led to a greater proliferation of connected things. Other areas with impact on the cost of deployment for IoT included:

- Decrease in the cost per CPU memory and storage makes possible the collection of big data, and its subsequent analytics
- Decreasing cost of megabytes increases the available investment dollars for large processing systems

2.2.4 LARGE IOT INVESTMENTS

IoT investments are increasing, predominantly the highest in the industrial markets. According to IDC, worldwide spending on the Internet of Things was forecast to reach $745 billion in 2019, an increase of 15.4 percent over spending in 2018. IDC expects worldwide IoT spending will maintain a double-digit annual growth rate throughout the 2017-2022 forecast period and surpass the $1 trillion mark in 2022.[32]

[31] *Ericsson Mobility Report.* November 2018.
[32] *IDC forecasts worldwide spending on the Internet of Things to Reach $745 billion in 2019,* led by the manufacturing, consumer, transportation, and utilities sectors, press release by IDC, 3 January 2019.

The largest IoT investments for 2019 are forecast by IDC as discrete manufacturing ($119 billion), process manufacturing ($78 billion), transportation ($71 billion), and utilities ($61 billion), largely focused on solutions that support manufacturing operations and production asset management. In transportation, more than half of IoT spending will go toward freight monitoring, followed by fleet management. The utilities industry spending will be dominated by smart grids for electricity, gas, and water. IDC expects that the industries with the fastest Compound Annual Growth Rate (CAGR) over the five-year forecast period (2017-2022) are insurance (17.1 percent), federal/central government (16.1 percent), and healthcare (15.4 percent).[33]

Consumer IoT spending will reach $108 billion in 2019, making it the second largest industry segment according to IDC research director Marcus Torchia. The leading consumer use cases will be related to the smart home (home automation and smart appliances in particular), personal wellness, and connected vehicle infotainment. Consumer IoT will be the fastest growing industry segment overall with a five-year CAGR of 17.8 percent.[34]

The IDC report further notes that IoT use cases that will see the greatest levels of investment in 2019 are driven by the industry spending leaders: manufacturing operations ($100 billion), production asset management ($44.2 billion), smart home ($44.1 billion), and freight monitoring ($41.7 billion). The IoT use cases that are expected to deliver the fastest spending growth over the 2017-2022 forecast period provide a picture of where other industries are making their IoT investments. These include airport facility automation (transportation), electric vehicle charging (utilities), agriculture field monitoring (resource), bedside telemetry (healthcare), and in-store contextualized marketing (retail).

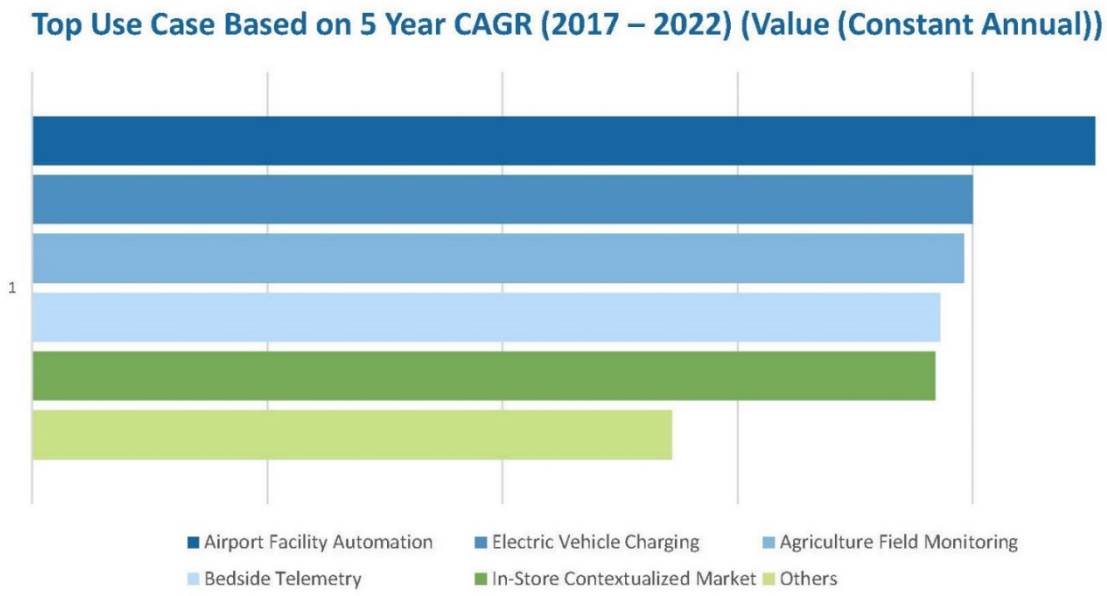


Figure 2.7. Top Use Cases Based on Spending Growth.[35]

[33] *IDC forecasts worldwide spending on the Internet of Things to Reach $745 billion in 2019,* led by the manufacturing, consumer, transportation, and utilities sectors, press release by IDC, 3 January 2019.
[34] *Ibid.*
[35] *Ibid.*

Investment in IoT in 2019 will occur in the following key categories according to the IDC report:

- Services - $258 billion for traditional IT and installation services and non-traditional device and operational services
- Hardware - $250 billion led by more than $200 billion in module/sensor purchases
- Software spending - $154 billion and the fastest growth (2017-2022) at CAGR 16.6 percent
- IoT connectivity - $83 billion

The United States and China will be the global leaders for IoT spending in 2019 at $194 billion and $182 billion respectively, followed by Japan ($65.4 billion), Germany ($35.5 billion), Korea ($25.7 billion), France ($25.6 billion) and the United Kingdom ($25.5 billion). The countries that will see the fastest IoT spending growth over the forecast period are all located in Latin America: Mexico (28.3 percent CAGR), Colombia (24.9 percent CAGR), and Chile (23.3 percent CAGR).[36]

Business Insider's survey data in 2019 shows that companies' plans to invest in IoT solutions are accelerating with 5G on the horizon and an uptick in IoT adoption.[37]

The continued growth of the IoT industry is going to be a transformative force across all organizations.

2.2.5 GLOBAL APPLICATION TRENDS AND USE CASES

IP video, in all its forms --Internet video, IP Video on Demand (VoD), video files exchanged through file sharing, video-streamed gaming, and video conferencing-- will continue to be in the range of 80 to 90 percent of total IP traffic. Globally, IP video traffic will be 82 percent of all IP traffic (both business and consumer) by 2022.[38] The implications of video growth are difficult to overstate. With video growth, Internet traffic is evolving from a relatively steady stream of traffic (characteristic of Peer-to-Peer [P2P] traffic) to a more dynamic traffic pattern.

New Internet-connected video surveillance cameras upload a constant video stream to the cloud for remote viewing. With a steady flow of video traffic from each camera, video surveillance is already having an effect on overall Internet traffic. It accounts for 2 percent of Internet video traffic today and will grow 7-fold to reach 3 percent by 2022. If such devices become mass market in the next five years, video cameras will generate a significantly higher volume of traffic, since Internet-enabled cameras can produce up to 300 GB per camera per month for full HD-resolution monitoring of high-activity areas.[39]

With new hardware available to individuals, and a growing body of content to consume, Virtual Reality (VR) and Augmented Reality (AR) are expected to continue a high growth trajectory through 2022. Traffic associated with VR and AR applications is poised to grow 12-fold over the next five years (65 percent CAGR) mainly stemming from the download of large virtual reality content files and applications, but a significant 'unknown' is the potential adoption of virtual reality streaming, which could further raise the prediction of high-growth.[40]

[36] *IDC forecasts worldwide spending on the Internet of Things to Reach $745 billion in 2019,* led by the manufacturing, consumer, transportation, and utilities sectors, press release by IDC, 3 January 2019

[37] *IoT Report: How Internet of Things Technology Growth is Reaching Mainstream Companies and Consumers*, Business Insider, Peter Newman. 28 Jan 2019.

[38] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.

[39] *Ibid.*

[40] *Ibid.*

The top three IoT projects in 2018 were Smart Cities (23 percent), Connected Industry (17 percent) and Connected Buildings (12 percent) according to IoT Analytics.[41] From a regional perspective, nearly half of the Smart City projects were in Europe (45 percent), while the Americas lead in the area of Connected Help with 55 percent, as shown in Figure 2.8.

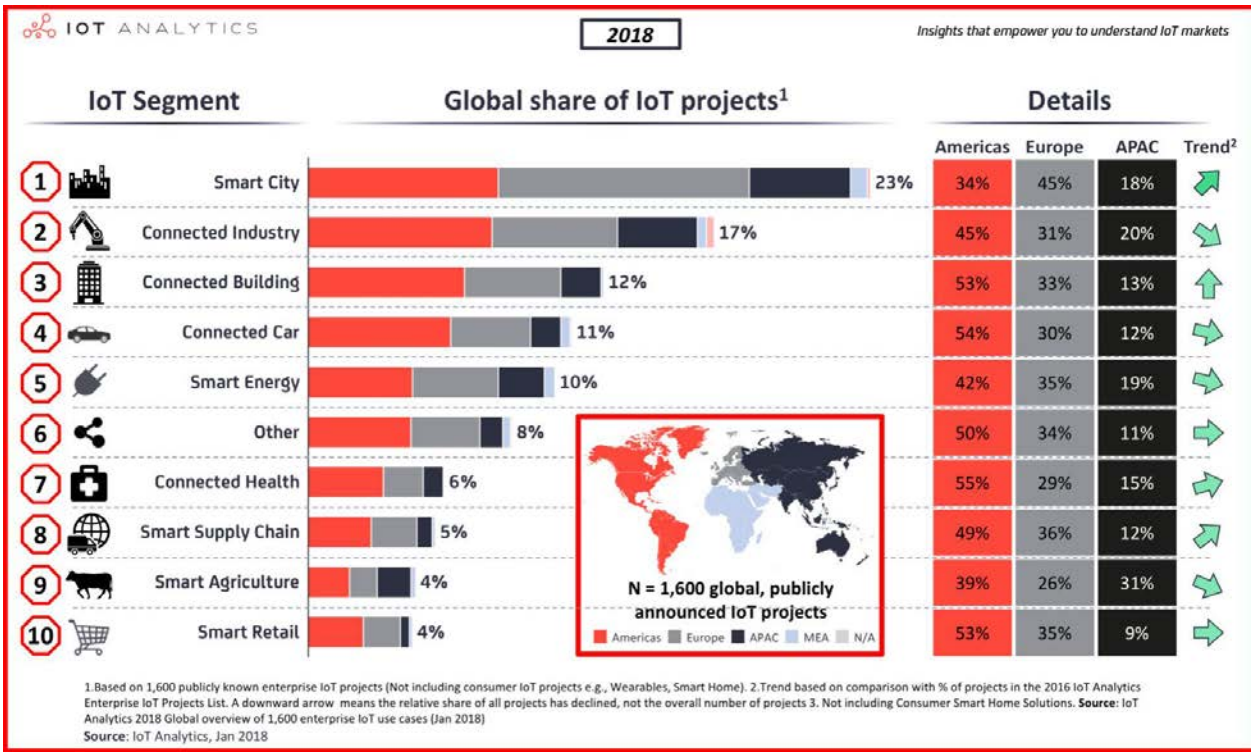


Figure 2.8. Global Share of IoT Projects by Segment – 2018.[42]

2.2.6 3GPP STANDARDS

3GPP has prepared for the IoT by developing standards beginning with Release 10. The lowest-cost cellular devices enabling M2M communications today are GPRS modems, which risk becoming obsolete as operator's sunset their Global System for Mobile communications (GSM) systems. High Speed Packet Access (HSPA) is also used for M2M communications, as is LTE, which has been optimized to efficiently communicate small bursts of information, making it particularly well suited for M2M. Low-cost GSM (through Enhanced Coverage GSM IoT [EC-GSM-IoT]) and LTE modem options in 3GPP Releases 10 through 13 reduce cost, improve communications range, and extend battery life.

In Release 13, the 3GPP standards body addressed study items for the IoT due to increasing demand from service providers and the influence of competition. This resulted in the creation of LTE-M and NB-IoT (previously explained in the Introduction of this paper). In Release 14, 3GPP specified how LTE technologies can operate for vehicle communications, including vehicle-to-vehicle and vehicle-to-infrastructure, leveraging LTE to 5G. These developments in Rel-14 provide LTE Highly-Reliable Low

[41] *The Top 10 IoT Segments in 2018 – based on 1,600 Real IoT Projects,* IoT Analytics, 28 February 2018.
[42] *Ibid.*

Latency Communications (HRLLC) for Critical IoT. Release 15 includes further IoT enhancements in LTE, including TDD support, higher spectral efficiency, and wake-up radio for 5G URLLC for Critical IoT.

The standards work has continued through Release 16, with ongoing developments to provide future-proof, secure and flexible technology for the IoT. Building on the foundation of the most pervasive mobile wireless technology – LTE – the new IoT standards are gaining momentum worldwide.

Service providers have announced the deployment of 68 cellular IoT networks worldwide as of July 2019: 24 using LTE Cat-M; 56 using NB-IoT; and including 12 operators with both IoT LTE-M and NB-IoT standards.[43] Both technologies are being deployed to complement each other across regions worldwide. Large-scale deployments, and the resulting high-volume of chipsets, are expected to continue to reduce chipset prices. As noted previously in 2.2.3, this is leading to further acceleration of the growth in cellular IoT connections.

For more information on the development of IoT standards requirements by 3GPP, reference the 5G Americas whitepaper published in October 2018, *Wireless Technology Evolution, Transition from 4G to 5G, 3GPP Releases 14 to 16.*

2.2.7 EMERGING NEW MOBILE BROADBAND TECHNOLOGY

NB-IoT and Cat-M1 will continue to coexist with the introduction of 5G networks. As new enabling IoT technologies are deployed, the number of connections and the traffic per connection over cellular networks will drive increasing traffic volumes, as network speeds continue to get faster.

Mobile speeds are increasing. Globally, the average mobile network connection speed in 2017 was 8.7 Mbps and that will more than triple to 28.5 Mbps by 2022 as shown in Table 2.1. Anecdotal evidence supports the idea that overall use increases when speed increases, although there is often a delay between the increase in speed and the increased use, which can range from a few months to several years. The reverse can also be true with the burstiness associated with the adoption of tablets and smartphones, where there is a delay in experiencing the speeds that the devices can support. The Cisco VNI Forecast relates application bit rates to the average speeds in each country. Many of the trends in the resulting traffic forecast can be seen in the speed forecast, such as the high growth rates for developing countries and regions relative to more developed areas.

[43] TeleGeography.July 2019.

Table 2.1. Global Mobile Network Speeds 2017-2022 by World Region.[44]

	2017	2018	2019	2020	2021	2022	CAGR 2017–2022
Global							
Global speed: All Connections	8.7	13.2	17.7	21.0	24.8	28.5	26.7%
Global speed: Smartphones	13.5	14.9	22.1	25.9	34.8	41.6	25.2%
Global speed: Tablets	22.6	24.5	25.9	32.9	40.5	57.2	20.4%
By Region							
Asia Pacific	10.6	14.3	18.0	21.7	25.3	28.8	22.1%
Latin America	4.9	8.0	11.2	13.0	15.3	17.7	29.6%
North America	16.3	21.6	27.0	31.9	36.9	42.0	20.9%
Western Europe	16.0	23.6	31.2	37.2	43.8	50.5	25.8%
Central and Eastern Europe	10.1	12.9	15.7	19.5	22.8	26.2	21.0%
Middle East and Africa	4.4	6.9	9.4	11.2	13.2	15.3	28.0%

Note: Current and historical speeds are based on data from Ookla's Speedtest. Forward projections for mobile data speeds are based on third-party forecasts for the relative proportions of 2G, 3G, 3.5G, 4G and 5G among mobile connections through 2022.

Source: Cisco VNI Mobile, 2019

IoT traffic volume is limited but increasing. To date, IoT has been characterized by a very large number of connections, small data volumes and, in some cases, stringent requirements on energy consumption. Typical uses are sensor, monitor or control data IoT applications. In many markets, ultra-low-end IoT applications with limited demands on throughput, such as sensors and monitoring, were the first services to be built on NB-IoT technology. Data traffic generated by such devices is generally low; the typical data packet for a sensor-based service is about 100–150 bytes, with a payload comprised of a device ID, time stamp and reported data values.[45] NB-IoT technology is capable of supporting data rates of 227 Kbps in uplink and 250 Kbps in downlink. Cat-M1 is a second technology designed and standardized for massive IoT applications and is capable of supporting data rates up to 1 Mbps in both the uplink and downlink. The traffic volume generated by massive IoT applications is a function of message size, message interval and number of devices deployed per square kilometer.

Beyond massive deployments of devices generating limited data volumes, there is an evolving range of IoT applications that have stringent requirements on availability, delay and reliability. Applications include traffic safety, automated vehicles, drones and industrial automation. These can generate many times more data traffic than massive IoT applications, depending on the specific use case. They could be based on LTE devices where, for example, an LTE Cat 4 device supports data rates of 150 Mbps in downlink and 50 Mbps in uplink; or have use case requirements that only future 5G network capabilities can meet. NB-IoT and Cat-M1 will continue to coexist with the introduction of 5G networks. The diagram in Figure 2.9 illustrates that, as new enabling IoT technologies are deployed, both the number of connections and the traffic per connection over cellular networks will drive increasing traffic volumes.

[44] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.
[45] *Ericsson Mobility Report.* November 2018.

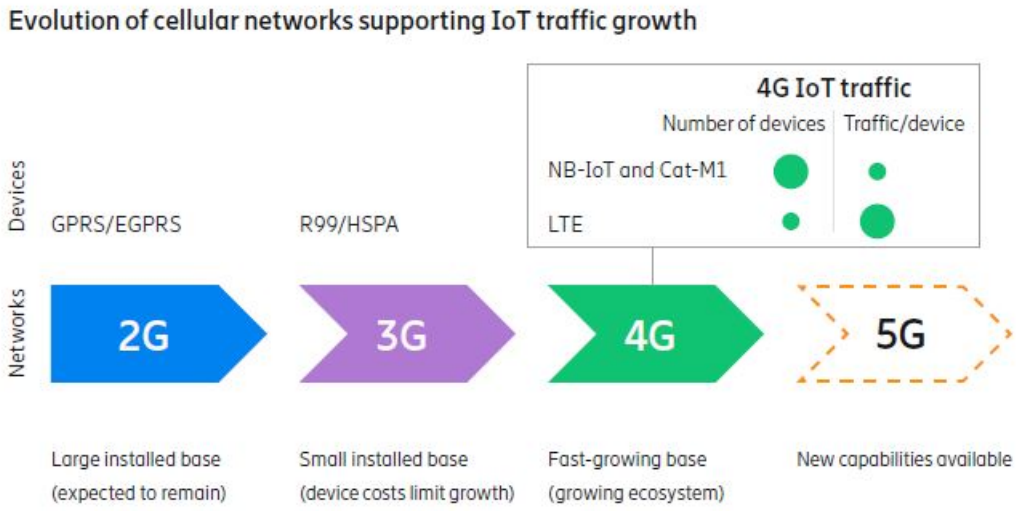


Figure 2.9. Evolution of Cellular IoT Technologies.[46]

2.2.8 GROWING IMPORTANCE OF AUTOMATION AND BIG DATA

There is a growing importance for automation, big data and other 'actionable knowledge' that is driving the IoT, the interconnection of various devices, machines, or appliances that generate data. IoT includes the connectivity of non-traditional devices, like vehicles, house appliances, smartphones, and smart gadgets that have electronic sensors and software embedded in their core systems. The underlying aim of IoT is not just to create data, but also to extract valuable insights and information from the data generated by these devices. Various industries, governments, and consumers to serve their needs can then effectively deploy these insights.

According to a McKinsey market analysis, around €23 billion will be generated in Germany in 2020 with the intelligent networking of machines and devices. In 2015, annual IoT sales in Germany were still under €10 billion, meaning the potential will more than double within five years. The most important fields of application for IoT are the digitalization of production (Industry 4.0) with a potential of just under €9 billion and networked vehicles at around €4 billion.[47]

2.2.9 ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Artificial Intelligence (AI) has been around in some form since the 1960s and today real use cases with valuable results are emerging as adoption steadily increases, particularly around machine learning (ML). According to research from McKinsey Digital, AI has caught on in the IoT in the past few years (2016-2018); AI and ML are being used in 60 percent of IoT activities.[48] Three major things have spurred the increase in the use of AI according to the report: the convergence of algorithmic advances, data proliferation, and

[46] *Ericsson Mobility Report.* November 2018.
[47] *The IoT as a Growth Driver,* McKinsey Digital. March 2018. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-iot-as-a-growth-driver
[48] *Ibid.*

tremendous increases in power and storage capabilities at a lower cost. Adoption of AI and ML expected to outpace other technologies as shown in Figure 2.10.[49]

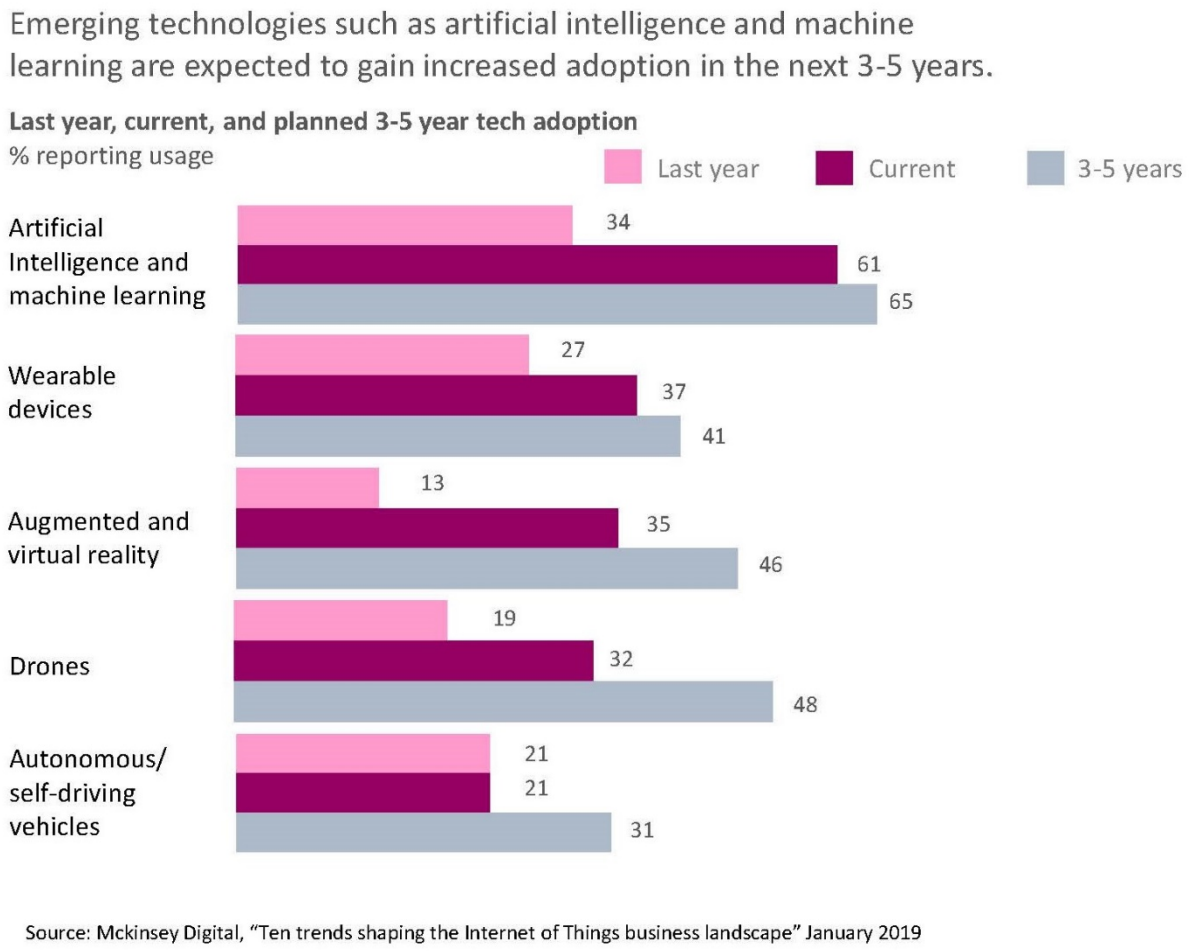


Figure 2.10. AI and ML Adoption in IoT 2017-2024.[50]

2.2.10 EDGE COMPUTING AND THE CLOUD

A common assumption among those new to IoT is that data need to be in the cloud or some similar central location in order to be analyzed. While this is true at times, as long as data transmission costs remain high, especially for remote industrial environments, performing some analytics at the 'edge' – therefore, adjacent to where the data originates - will become an option. In many industrial sectors with mobile and/or remote assets (such as oil and gas, aviation, and transportation), shifting some analytics intelligence to the edge may be more cost effective. Autonomous vehicles face a similar challenge; even with better data-transport technologies such as 5G, response times for rapidly moving vehicles may make an edge-based solution more relevant. For the most part, the debate about whether to store data and analytics at the edge or

[49] *Ten trends shaping the Internet of Things business landscape*, McKinsey Digital. January 2019.
[50] *Ibid.*

centrally on the cloud as the IoT host environment hinges on which is decreasing faster: the cost (both are declining in cost) and latency of data transmission or the cost of "smarter" edge equipment.[51]

Although not often listed as a key market driver for the IoT, Edge Computing and the Cloud might be considered as a wild card due to network architecture transformation. Edge networking continues to gain more intelligence and capacity to support evolving network demands and superior network experiences. Increasingly, global service providers are making networking investments and architectural transformations to bolster the capabilities of the network edge.[52]

Based on analysis by Cisco, 33 percent of global service provider network capacity will be within a metro network by 2022 (up from 27 percent in 2017). Comparatively, 24 percent of global service provider network capacity will be in regional backbones by 2022 (down from 25 percent in 2017) and 43 percent of global service provider network capacity will be in cross-country backbones by 2022 (down from 48 percent in 2017).[53]

Edge computing plays a major role in IoT functionality. The low latency and reliability edge computing provides are requisites for most IoT use cases. When automated vehicles communicate with each other about hazards on the road, they need the low latency provided by the nearby edge network to spread information fast enough to avoid crashes. Additionally, a smart factory can't afford to stop production because the larger network has gone down. If it takes advantage of edge computing, the interconnected system of machines can keep running.

According to IDC, IoT deployments are fueling aggressive investments in infrastructure for new compute, storage and networking technologies at the edge. Deploying edge infrastructure, in turn, will drive many Greenfield implementations resulting in growth of the IoT market.[54]

2.2.11 MORE ADVANCED USE CASES ACROSS VERTICAL DOMAINS

M2M applications across many industries are accelerating IoT growth. Industrial IoT, Smart Cities, Enterprise IoT and Consumer IoT including Smart Homes, Wearables, and Connected Car are all contributing to market expansion and are explained more fully in sections 2.5 to 2.8. More advanced IoT use cases requiring enhanced network capabilities are emerging as the IoT application market is widening. Examples of such capabilities are: support for optimized voice quality; more accurate device positioning; and support for device mobility at high speed.

[51] *The IoT as a Growth Driver*, McKinsey Digital. March 2018. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-iot-as-a-growth-driver
[52] *A Driving Force Behind the EDGE: IoT,* SDxCentral. April 2019.
[53] *Cisco Visual Networking Index*: Forecast and Trends, 2017-2022, white paper. Updated 27 February 2019.
[54] *Market Forecast,* Doc # US44154318, Tech Supplier. July 2018.

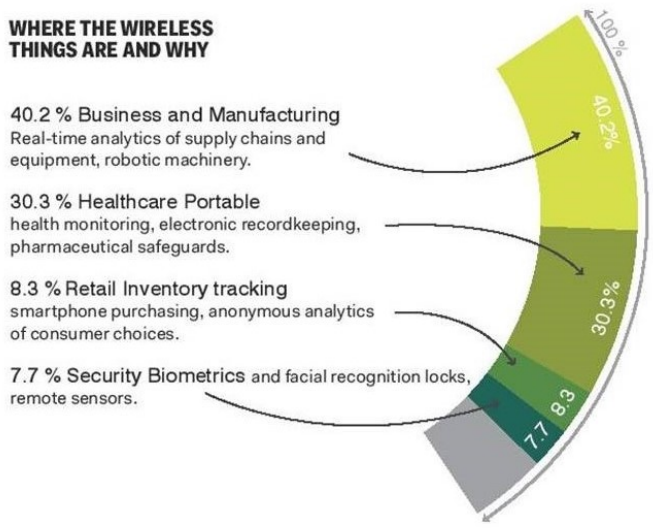


Figure 2.11. Key Verticals for IoT.[55]

Connected home applications, such as home automation, home security and video surveillance, connected white goods, and tracking applications, will represent 48 percent, or nearly half, of the total M2M connections by 2022, showing the pervasiveness of M2M in our lives, as shown in Figure 2.11.

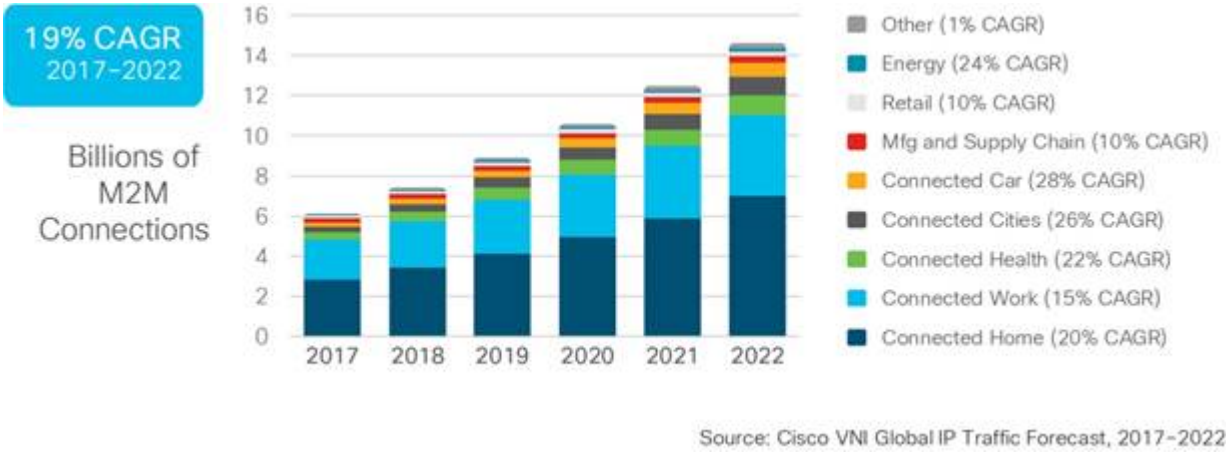


Figure 2.12. Global M2M Connection Growth by Industries.[56]

Connected car, with applications such as fleet management, in-vehicle entertainment and Internet access, roadside assistance, vehicle diagnostics, navigation, and autonomous driving, will be the fastest-growing industry segment, at a 28 percent CAGR. Connected cities applications will have the second-fastest growth, at a 26 percent CAGR each as noted in Figure 2.12.

[55] *Digitize or Die*, Nicolas Windpassinger.
[56] *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.

2.2.12 SECURITY ASSURANCE

In addition to traffic growth ramifications, IoT is also a catalyst for fixed/mobile convergence, network innovations and comprehensive network security improvements. Security assurance is an essential motivator for vertical industries as well as consumers – and risk managers now consider cyber risk to be the biggest threat to their businesses. According to 2017 McKinsey survey, 75 percent of experts consider cybersecurity to be a top priority.[57] As growth in most industries depends on new technology, such as artificial intelligence, advanced analytics, and the Internet of Things (IoT) bringing all kinds of benefits, it also exposes companies and their customers to new kinds of cyber risk, arriving in new ways.

By 2020, the IoT may comprise as many as 30 billion devices, many of them outside corporate control. Already, smart cars, smart homes, and smart apparel are prone to malware that can conscript them for distributed denial-of-service attacks. By 2020, 46 percent of all Internet connections will be machine-to-machine, without human operators, and this number will keep growing. And of course, billions of chips have been shown to be vulnerable to Meltdown and Spectre attacks, weaknesses that must be addressed.

McKinsey suggests that companies need to embrace and adopt automation, big data solutions and artificial intelligence to cope with the ever-increasing number of alerts and incidents. To be effective, though, the organization needs a company-wide governance structure, built on a strong cyber risk culture. Governance of IT, Operating Technology (OT), the IoT, and products should be consolidated into one operating model, and the entire business system should be covered, including third parties.

Fortunately, 5G is not just about faster, bigger or better. It's about enabling a diverse new set of services and use cases affecting nearly every aspect of our lives. But to live up to their potential, 5G-enabled applications must be delivered securely.

For example, 5G will enable Massive Internet of Things (MIoT) applications such as the traffic sensors and Vehicle-to-Infrastructure (V2I) services that are the foundation for smart cities. It's critical that hackers can't access that data, hijack IoT devices or disrupt the services with Distributed Denial of Service (DDoS) attacks.

Fortunately, security has been a top architectural priority with all previous mobile generations. 3GPP Release 8 added a variety of advanced security/authentication mechanisms via nodes such as the services capability server, while Release 11 provided additional capabilities to enable secure access to the core network. These and other 4G-era additions are noteworthy because LTE is the foundation for 5G, including its security mechanisms. Much of the information in this section has been condensed from the 5G Americas whitepaper, *The Evolution of Security in 5G*, published in October 2018.[58]

The mobile wireless industries longstanding emphasis on security has been a strong market differentiator against many other wireless technologies which have network architectures that are inherently more vulnerable. Even mobile's use of licensed spectrum provides a powerful additional layer of protection against eavesdropping on data, voice and video traffic.

With 5G, mobile takes that security focus to another level with a wide variety of new, advanced safeguards. The 5G Americas security white paper describes those safeguards in depth, as well as the vulnerabilities and attack vectors that they're designed to mitigate. It also explores how 5G differs from 4G and 3G in

[57] *A new posture for cybersecurity in a networked world,* article by Thomas Poppensieker and Rolf Riemenschnitter, McKinsey. March 2018.
[58] *The Evolution of Security in 5G,* 5G Americas Whitepaper. October 2018.

terms of radio and core network architectures, and how those differences affect the security mechanisms available to mobile operators, their business partners and their customers.

For example, 5G is the first mobile architecture designed to support multiple, specific use cases, each with their own unique cybersecurity requirements. In the enterprise IT world, network segmentation is a common, proven way to mitigate security risks. 5G introduces the concept of network slicing, which provides mobile operators with segmentation capabilities that weren't possible with previous generations.

With the IoT, security challenges move from a company's traditional IT infrastructure into its connected products in the field. And these challenges remain an issue through the entire product life cycle, long after products have been sold. What's more, industrial IoT, or Industry 4.0, means that security becomes pervasive in production as well. Cyberthreats in the world of IoT can have consequences beyond compromised customer privacy. Critical equipment, such as pacemakers and entire manufacturing plants, is now vulnerable—meaning that customer health and a company's total production capability are at risk.

The sheer number of cybersecurity attack vectors increases dramatically as ever more "things" are connected. Earlier, a large corporate network might have somewhere between 50,000 and 500,000 endpoints; with the IoT, we are talking about millions or tens of millions of endpoints. Unfortunately, many of these consist of legacy devices with inadequate security, or no security at all. This added complexity makes the IoT a more difficult security environment for companies to manage. Those that succeed, though, could use strong cybersecurity to differentiate themselves in many industries.

Cybersecurity is or should be top of mind for virtually every CXO involved with the IoT; additional research from McKinsey indicates that almost 50 percent have been attacked with 25 percent experiencing high or severe damage as a result.[59] That said, even companies that have been attacked are for the most part not significantly curtailing their IoT activities. In short, cybersecurity is a big concern, but not a barrier to IoT adoption in most cases. Companies doing IoT at scale view it as a strategic imperative, and while they may change policy and invest more in cybersecurity, they are not ratcheting back IoT activities.[60]

[59] *The IoT as a Growth Driver*, McKinsey Digital. March 2018. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-iot-as-a-growth-driver
[60] *Ibid.*

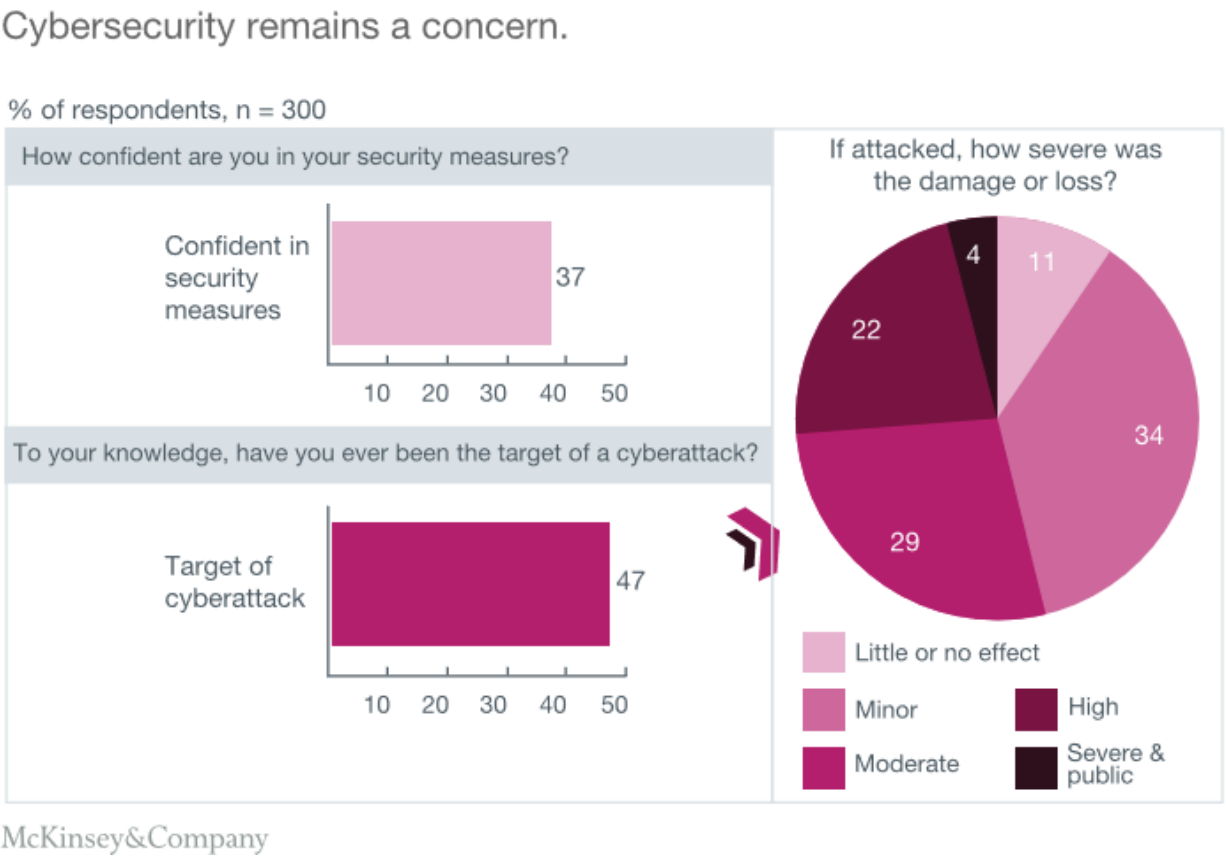


Figure 2.13. Security Concern for IoT.[61]

2.2.13 IPV6

IPv6 adoption enables IoT connectivity. Transition from an IPv4 environment to an IPv6 environment is making excellent progress, with increases in IPv6 device capabilities, content enablement, and operators implementing IPv6 in their networks. These developments are particularly important because Asia, Europe, North America, and Latin America have already exhausted their IPv4 allotments, and Africa is expected to exhaust its allotment in 2019. Building on the Cisco VNI IPv6-capable devices analysis:[62]

> The Cisco forecast estimates that globally there will be nearly 18.3 billion IPv6-capable fixed and mobile devices by 2022, up from nearly 6 billion in 2017, a CAGR of 26 percent. In terms of percentages, 64 percent of all fixed and mobile networked devices will be IPv6-capable by 2022, up from 32 percent in 2017. This estimate is based on the capability of the device and the network connection to support IPv6 and is not a projection of active IPv6 connections. Mobile-device IPv6 capability is assessed based on Operating System (OS) support of IPv6 and estimations of the types of mobile network infrastructure to which the device can connect (for example, later than 3G).

Looking to 2022, if 60 percent of IPv6-capable devices are actively connected to an IPv6 network, the Cisco forecast estimates that globally IPv6 traffic would amount to 132 EB per month, or 38 percent of total

[61] *The IoT as a Growth Driver*, McKinsey Digital. March 2018. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-iot-as-a-growth-driver
[62] Cisco *Visual Networking Index: Forecast and Trends, 2017-2022*, white paper. Updated 27 February 2019.

Internet traffic. This initial estimation of potential IPv6 traffic is based on the assumptions that IPv6 device capability, IPv6 content enablement, and IPv6 network deployment will keep pace with current trends and may even accelerate during the forecast period. Considering the interdependence of these variables, forecast assumptions could be subject to refinement.

Content providers are also moving to increase the IPv6 enablement of their sites and services. According to Cisco® IPv6 labs, by 2022 the content available over IPv6 will be about 51 percent. There can be, however, variation depending on the popularity of websites across regions and countries. In addition, specific country initiatives and content-provider deployments have positively affected local IPv6 content reachability. Overall, the likelihood that a significant portion of Internet traffic will be generated over IPv6 networks holds considerable opportunity for network operators, content providers, and end users seeking to gain the scalability and performance benefits of IPv6 and enable the Internet of Things (IoT).

2.2.14 OPEN SOURCE

The term “open source” refers to something people can share, modify and use via an openly available design. Open source, as applied to software, permits sharing via inspection, copying, learning, altering or distribution.

In order to provide massive amounts of bandwidth to a massive number of devices, there is a need to transform the network to be able to scale up and be agile while reducing cost. Network disaggregation with separation of user and control plane, separating out the network operating system from the underlying hardware, and use of general-purpose processing platforms is the key to creating networks that are massively scalable, agile and inexpensive. Open source will have an increasingly important role in future 5G networks, and 5G Americas published a whitepaper, *The Status of Open Source for 5G* in February 2019, from which the information in this section is largely attributed.[63]

Leveraging open source is important for enabling a high-performance, flexible 5G user plane. There are various open-source networking initiatives—such as Data Plane Development Kit (DPDK), Vector Packet Processing (VPP), Fast Data Input/Output Project (FD.io), Mobile Central Office Re-architected as a Datacenter (M-CORD), National Ground Intelligence Center (NGIC) and Open Virtualized multilayer Switch (Open vSwitch or OVS) —that provide the necessary optimizations, bringing in the ability of the user plane to scale and handle increased throughput necessary for 5G use cases and services.

5G brings a diverse set of requirements and use cases, requiring an entirely new RAN architecture that is flexible, modular and supports open interfaces. The new RAN architecture needs to be operationally efficient and able to dynamically adapt to various, diverse requirements of 5G. The key to effectively implementing the new RAN architecture with flexible splits and efficient fronthaul is the openness in its specification and implementation.

Such proprietary software and interfaces are often tied to the underlying hardware, which is a significant roadblock for openness. Enabling multi-vendor, best-of-breed flexibility in the RAN requires a move away from proprietary hardware to off-the-shelf, general-purpose processing platforms. Adoption of such commodity network hardware will require a reference design with standardized interfaces.

There are several industry forums such as O-RAN, (Open - Radio Access Network) / Telecom Infra Project (TIP), which are focusing on decoupling the RAN control plane from the user plane, building a modular RAN software stack that uses commodity hardware and publishing open north- and south-bound

[63] *The Status of Open Source for 5G*, 5G Americas white paper, February 2019.

interfaces. The challenge has been the lack of openness in the RAN architecture, which is being addressed with the specifications being defined by the ORAN group.

The core network is a critical component, so it needs to be robust, highly resilient and high performance. 3GPP's Service-Based Architecture (SBA), has standardized the Network Functions (NFs), their procedures and the inclusion of NF sub-modules. Standards also define the APIs to be used by providing data model, protocol and format. However, there still exists lots of innovation and ongoing research in the open source community and within vendors to address specific problem areas or new ways of implementing specific interfaces.

The 5G industry vision goes beyond mobile, which is why 5G convergence in next-gen fixed and mobile cores will need to be addressed. Some additional areas to be considered as applicable for open source in the domain of management and control are:

- **Orchestration** of network services to provide expected agility in the telecom networks requires a way to define these services down to their atomic nature, physical and virtual resources
- **Network Automation** enables automation by programmatically configuring and provisioning network connections. The key to efficient network automation is maximum openness without which integration is not possible. Without standard bodies focused on network automation, there is a need for adopting an open source approach.
- **Analytics** will also be key to network automation, as it allows closed-loop feedback for effective service assurance. This will allow the network to self-heal, self-optimize and self-organize, bringing operational efficiency to network management. With the growing number of devices and data explosion in 5G networks, real-time analytics of disparate data will be necessary
- **DevOps** is a software engineering culture and practice that aims at unifying software development and software operation. The cloud-native approach is fundamental to 5G network functions/services and allows vendors and service providers to impose DevOps methods to automate the process of building, validating and deploying workloads into NFV environments. This enables service agility. Any open source efforts in this DevOps area are applicable for 5G
- **Testing Tools -** 5G Next Generation Core's (NGC) use of Open Application Programming Interface (API) will allow many existing test tools to be leveraged

There is an increased focus on security in 5G and open source software due to continuous evolving threat landscape and dynamically changing critical infrastructure that will carry massive amount of traffic to service multiple deployment options and use cases. As new features and functionalities are constantly added to the original open source software stack, all software components along with platform code must have a clean baseline for continuous security and compliance validation. Operators, solution providers and open source communities will have to manage accurate inventories of open source software dependencies to mitigate the risk of code injection. This is a risk to any software; companies utilizing strong auditing and sourcing processing can minimize their exposure to attacks. Processes will have to be put in place to receive and manage notifications concerning discovered vulnerabilities or available patches from the community supporting the open source.[64]

A key goal of the collaborative telecom industry effort is to encourage implementation and realization of the full potential of 5G working in tandem with the open source ecosystem. Security tools for 5G and open source environments are still evolving. Hence, a holistic approach to developing security in a 5G environment is required. Security controls need to be both horizontally and vertically distributed across the 5G environment to help enhance the end-to-end security posture. 5G transformation will occur; however,

[64] *CSRIC Network Reliability and Security Risk Reduction Report,* https://www.fcc.gov/files/csric6wg3sept18report5gdocx-0.

the level of its success will depend upon realizing the full potential of open source and making it securely deployable and operational.

Open source has been key to many recent advancements in high-performance and flexible packet processing, which is key to 5G use case and services. Analysis performed in the referenced 5G Americas paper suggests infrastructure (therefore, programmable forwarding plane, general purpose processing platforms), management and control (therefore, automation, orchestration, analytics, testing) are the key areas that, if open sourced, could benefit mobile operators the most in their 5G deployments.

5G system architecture gives mobile operators more openness than previous generations. Operators and OEMs may leverage open source principles in order to stay or become competitive in the marketplace.

5G will enable the IoT, with the ability to connect more than 10's of billions of sensors in the next decade. This level of scale may be supported by open source frameworks and platforms within that infrastructure. Open source technologies support rapid innovation through several advantageous characteristics. Typically, free and generally easy to download, install and launch allowing easy experimentation with new technologies. It also allows 'permission-less' innovation, easing concerns over Intellectual Property Rights. It also permits innovation by integration, meaning developers create new systems by combining freely available open source components. Open source software tends to promote innovation faster than proprietary solutions because they draw contributions from a large community of developers.

Interoperability is one of the best possible ways to have a new technology achieve rapid adoption by combining open standards with robust open source implementation thus reducing not only time to market but also cost of entering the market (time is money). Another advantage for open source in IoT is that the businesses are not locked to a proprietary vendor. Any cost of switching vendors is no longer an issue.

2.3 MARKET ANALYSIS FOR IOT

There is an emerging trend toward communications service providers deploying one IoT network that supports both Cat-M1 and NB-IoT technologies. This enables them to address the diverse and evolving requirements across a wide range of use cases in different verticals, such as utilities, smart cities, logistics, agriculture, manufacturing and wearables.

Massive IoT cellular technologies such as NB-IoT and Cat-M1 are taking off and driving growth in the number of cellular IoT connections worldwide. Table 2.1 shows Ericsson's IoT connections forecast, where the cellular IoT connections category is part of the wide-area IoT segment. Of the 4.1 billion cellular IoT connections forecast for 2024, North East Asia is anticipated to account for 2.7 billion – a figure reflecting both the ambitions and size of the cellular IoT market in this region.[65] These complementary technologies support diverse low-power wide-area (LPWA) use cases over the same underlying LTE network.

[65] *Ericsson Mobility Report*. November 2018.

Table 2.2. IoT Connections (Billion).[66]

IoT connections (billion)

IoT	2018	2024	CAGR
Wide-area IoT	1.1	4.5	27%
- Cellular IoT[2]	1.0	4.1	27%
Short-range IoT	7.5	17.8	15%
Total	8.6	22.3	17%

The market of Internet of things is accelerating unexpectedly and it growth is forecasted to reflect the increase in consumer expending and driven by the increase in IoT devices as detailed in the IoT Analytics report and illustrated in Figure 2.14.[67]

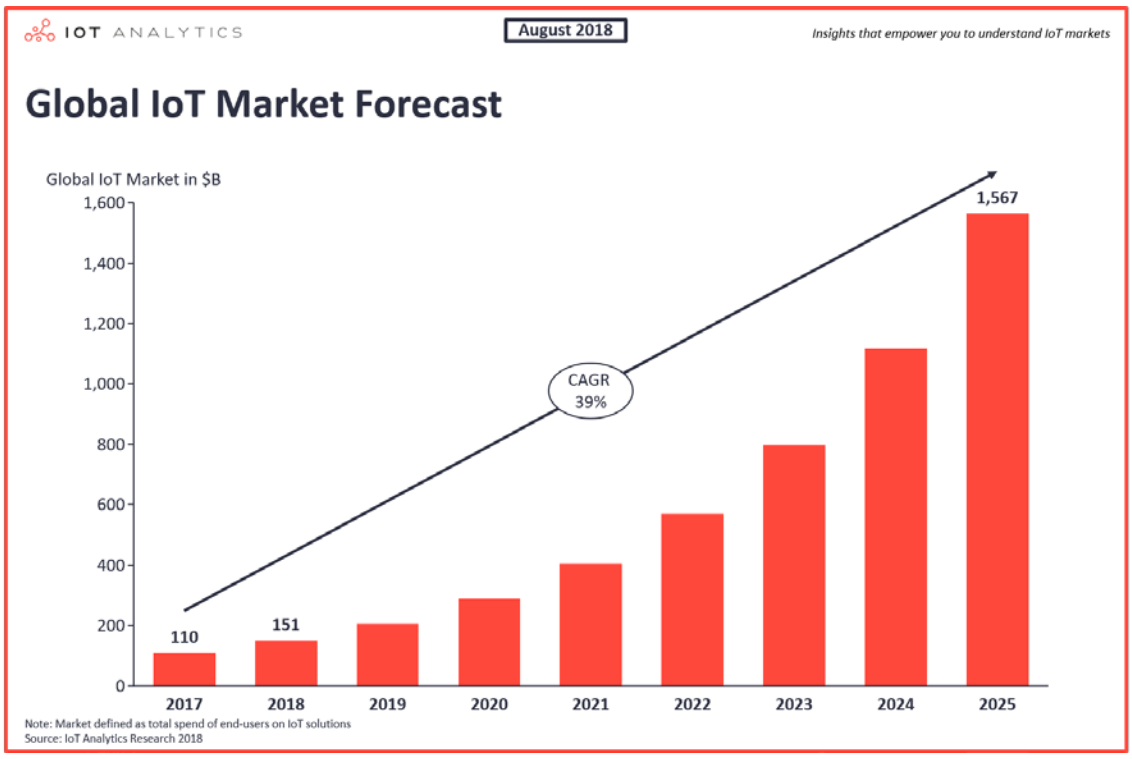


Figure 2.14. Global IoT Market Forecast.[68]

While the growth in IoT is linked with the increase of IoT connected devices as shown in Figure 2.15, the 5G technology promises a new era of connectivity through its massive bandwidth and extremely low latency. In addition, 5G is expected to have a driving effect on high-end IoT applications linked to smart

[66] *Ericsson Mobility Report*. November 2018.
[67] *State of the IoT & Short-Term Outlook 2018*, report by IoT Analytics, 2018.
[68] *Ibid.*

cities and industrial IoT, together with robotics, automation, and Artificial Intelligence with Machine Learning (AI/ML).

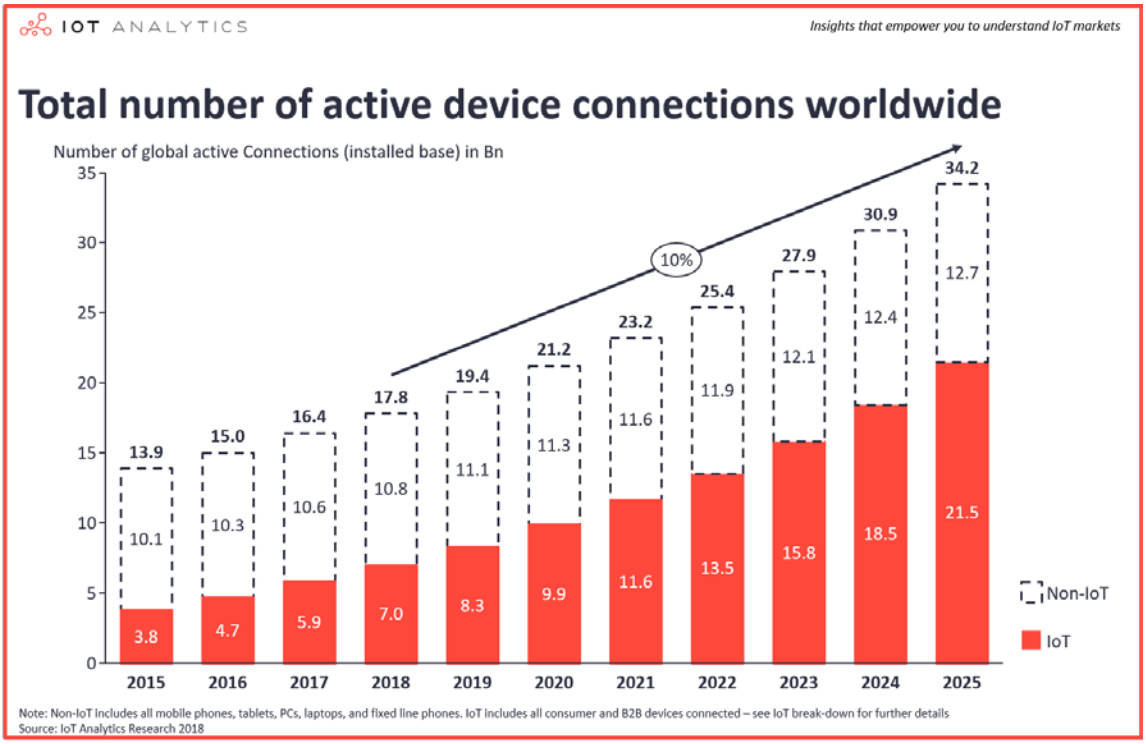


Figure 2.15. Projected Global Growth in IoT Devices.[69]

2.4 IOT PLATFORMS

Differentiating the type of IoT use cases or platform sectors is not so much based upon the 'who' or 'where' but on the requirements for the technology implemented. In other words, the difference between consumer IoT, Enterprise IoT and Industrial IoT (IIoT) is not that the device is used in the home, in an office or in a factory, but on the requirements established for the IoT connection. A consumer IoT device may have the same functionality as an IIoT device, and still not be considered an industrial product; for example, a consumer and an industrial activity tracker both collect and monitor heart rate information. But the industrial tracker incorporates additional design parameters that its consumer counterpart may not have. The parameters that differentiate IoT from industrial IoT include:[70]

[69] *State of the IoT & Short-Term Outlook 2018*, report by IoT Analytics. 2018.
[70] *IoT vs. Industrial IoT: 10 Differences that Matter,* Benson Chan, IoT for All, 14 December 2017.

- Security
- Interoperability
- Scalability
- Precision and Accuracy
- Programmability
- Low latency
- Reliability
- Resilience
- Automation
- Serviceability

Of these parameters, security is critical for all IoT solutions, but industrial IoT solutions require even more robust measures than either consumer or enterprise IoT, with the exception of Connected Car. For example, a takedown of the electrical grid affects economic activity for millions of people and jeopardizes national security. IIoT solutions employ a variety of advanced security measures, from secure and resilient system architectures, specialized chipsets, encryption and authentication, threat detection, to management processes.

Each of these key IoT categories are explained in sections 2.5 to 2.7: Industrial IoT; Enterprise IoT; and Consumer IoT.

2.5 INDUSTRIAL IOT

Industrial automation or Industrial IoT (IIoT) initially referred to an industrial framework whereby a large number of devices or machines are connected and synchronized through the use of software tools and third platform technology in an M2M and IoT context. Later the term Industry 4.0 evolved as well. Today, it is mainly distinguished from consumer IoT and includes major industries such as oil and gas, transportation, manufacturing, healthcare and energy. However, there are many other use cases in government, such as smart cities, and agriculture. There is also delineation (not always clear delineation) with Enterprise IoT and as previously explained this differentiation is not who or where, but the requirements particularly regarding strict security measures needed for IIoT. Thus, IIoT is another potential area where 5G IoT can help improve production efficiency and safety. It involves connecting industrial automation sensors, devices and equipment with cloud-based systems to gain business value. The Industrial IoT market size is projected to be $270 billion by 2024.[71]

Cellular network capabilities are evolving from the support of massive IoT (MIoT) to extreme low-latency IoT applications – meeting the requirements of IIoT. Currently, most use cases on manufacturing sites are based on wired connections. However, as the evolving cellular capabilities are challenging industrial ethernet solutions, cables will in many cases become redundant, introducing opportunities for more flexible production and expanded digital operations.[72]

Smart wireless manufacturing is currently being introduced. Eventually the legacy installed fixed network technologies will be incapable of effectively managing the use case requirements in advanced manufacturing. The opportunity for service providers and their network technology partners is to create a new market for smart wireless manufacturing by bridging the perceived value gap, addressing pain points

[71] *Industrial Control and Factory Automation Market - Global Forecast to 2024*, MarketsandMarkets Analysis, April 2019.
[72] *Realizing smart manufacturing through IoT,* Ericsson Mobility Report. June 2018.

for manufacturers/industry to offset switching costs and proving cellular IoT's business and practical value, and build horizontal and scalable solutions to address cost, deployment and spectrum issues.

Key sectors that will drive engagement for IIoT include:

- processes requiring mobility, such as shop floors with automated vehicles, and assembly warehouses that need secure and precise management as well as tracking of traffic, data flows and assets
- low-volume and high-variance manufacturing cases, where wireless machine line configuration is simple and flexible compared to cabled machine lines
- processes that cannot be monitored and controlled via cables but require wireless, real-time critical data transmission and a stable, deterministic network performance (bandwidth and latency) to operate
- processes susceptible to human error, or advanced manufacturing that requires tacit knowledge and skills transfer, where digital tools will be widespread to mitigate for errors and encourage faster learning
- processes where coordination of factories, resources and components is time-sensitive or crucial for the result (for example, product quality and timely delivery)

The choice of connectivity will determine the quality and flexibility of a manufacturer's digital foundation, as well as the possibilities and ultimately the value it will bring to their operations. It affects which equipment and operations can be connected, how many assets and processes can run simultaneously, and how well it scales beyond one geographical site. Manufacturing companies that exploit the full value of using cellular networks' global, wide-area capabilities beyond a single manufacturing site will also explore increased internal and external collaboration, creating tighter value networks with partners and other stakeholders. With the expected growth in demand for automated, customized, remote, and even mobile production, the need for enhanced wireless network capabilities will only increase.

According to Forrester Research in a 2018 study, industrial products lead all industries in IoT adoption at 45 percent with an additional 22 percent planning to adopt IIoT in 2019.[73] This is further supported by Accenture's forecast that the Industrial IoT will reach $123 billion in 2021, attaining a CAGR of 7.3 percent through 2020.[74] Accenture also forecasts that IIoT can add as much as $14.2 trillion to the global economy by 2030.[75]

[73] *10 Charts that will challenge your perspective of IoT's growth*, Forbes. 6 June 2018.
[74] *Ibid.*
[75] *Ibid.*

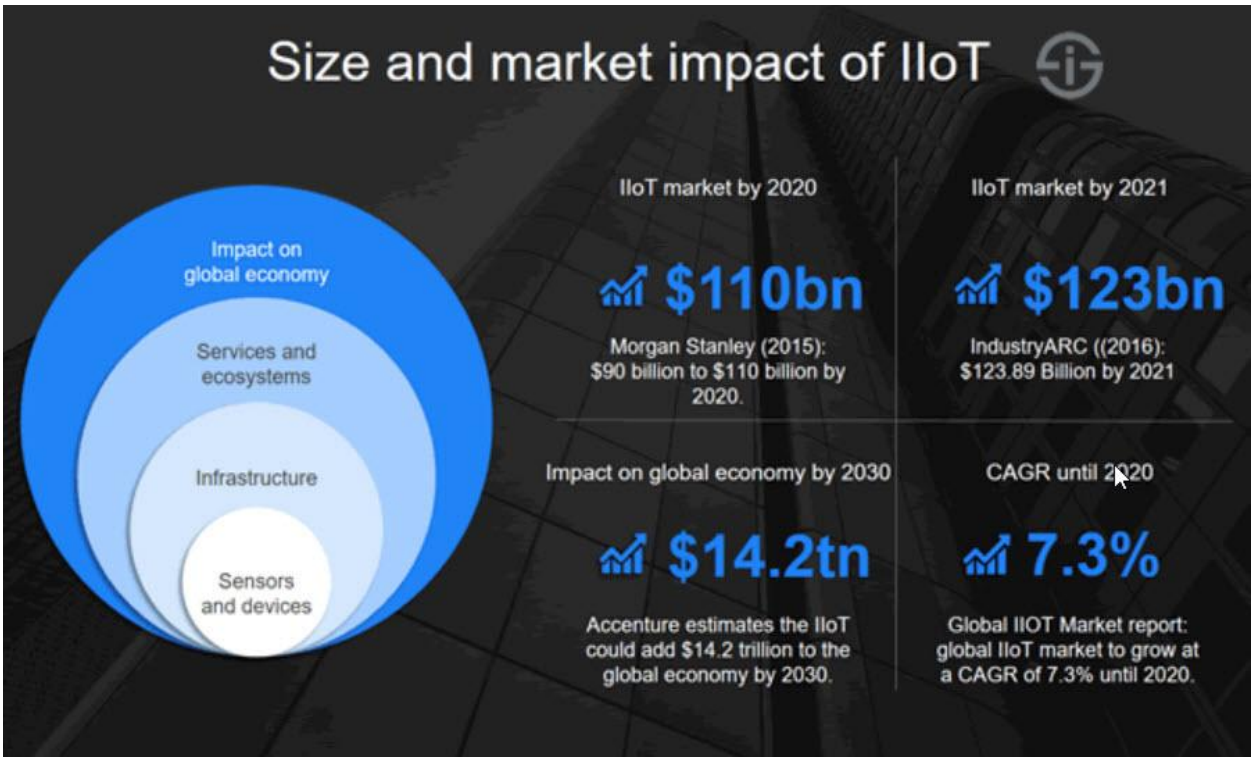


Figure 2.16. Size and Market Impact of IIoT.[76]

Entire Industry 4.0 use cases based on robotics and slow-moving robots use techniques such as imaging, sensing and Extended Reality (XR) to increase production efficiencies. These highly flexible factories and plants could use 5G IoT bands for the purpose of achieving very high wireless data throughput or use of a non-public (private or dedicated) network for other IoT on devices. These types of devices would operate in environments where radio frequency propagation characteristics may not be affected drastically, and are not be too restricted by power consumption, multi-antenna locations, or beamforming performance. 5G IoT cellular networks could provide highly reliable connectivity, enabling factories to become less dependent on wires and more flexible, particularly using edge computing. By capturing information in real-time and enabling remote control of machinery and automatically adapt to shop-floor events, increasing efficiency can be achieved by manufacturers for their customers.

GE found that IIoT applications are relied upon by 64 percent of power and energy (utility) companies to succeed with their digital transformation initiatives.[77]

[76] *The Industrial Internet of Things (IIoT): The Business Guide to Industrial IoT*, I-Scoop. Source: *Morgan Stanley, IndustryARC, Accenture and Research and Markets*. 2018.
[77] *The Industrial Internet of Things (IIoT): The Business Guide to Industrial IoT*, I-Scoop. Source: *GE Digital Industrial Evolution Index Executive Summary*. October 2017 (PEF, 67 pp, no opt-in).

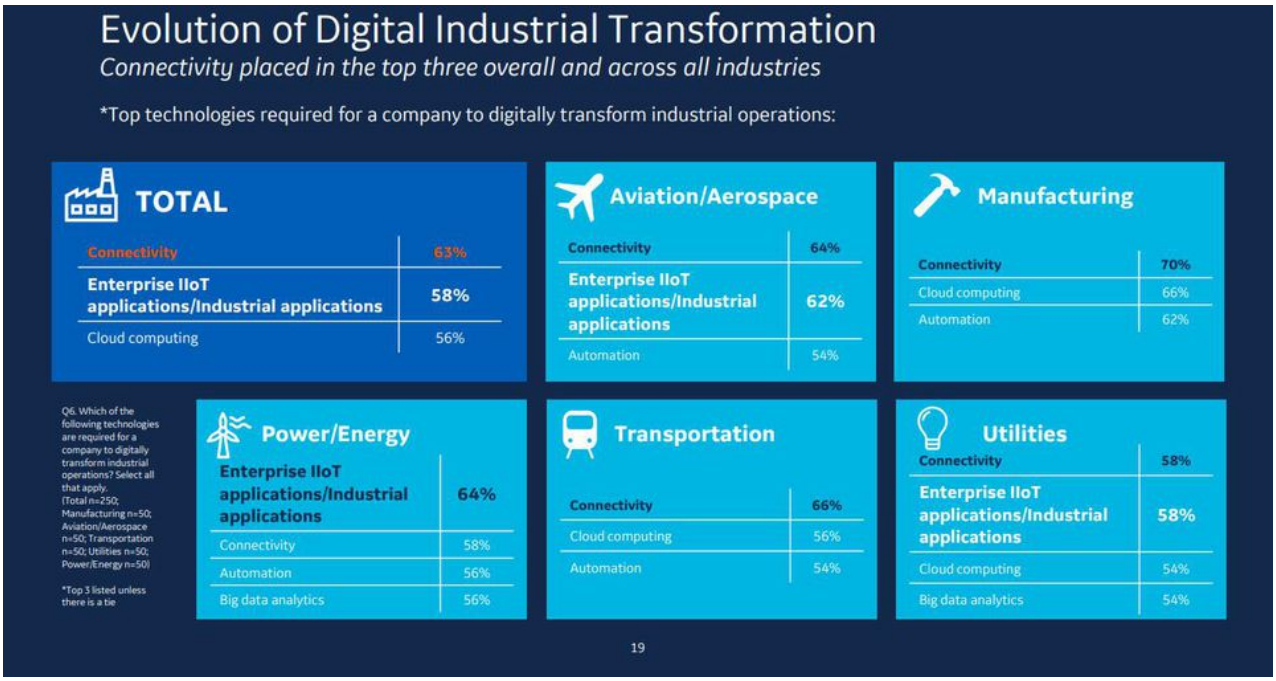


Figure 2.17. Evolution of Digital Industrial Transformation.[78]

Factory automation is one use case where products are produced, assembled, tested or packed in many discrete steps (automotive, general consumer electronic, goods production). In-time deliveries of messages and high reliability are very important to avoid interruptions in the manufacturing process. Redundancy, security and functional safety are also very important for factory automation. Typically, every manufacturing step involves many sensors and actuators controlled by a single Programmable Logical Controller (PLC). Many of these PLCs use wired connections which are often stressed by repeated movements and/or rotations and other harsh conditions. More and more devices, especially sensor and actuator nodes, will be connected using 5G IoT technology to improve productivity and increase availability compared to wired sensors/actuators at difficult locations. When connection density is very high, 5G mmWave is well-suited to provide needed capacity and minimize the end-to-end latency for the factory network. This is diagrammed in Figure 2.18 based on 3GPP specifications.

[78] *The Industrial Internet of Things (IIoT): The Business Guide to Industrial IoT*, I-Scoop. Source: *GE Digital Industrial Evolution Index Executive Summary*. October 2017 (PEF, 67 pp, no opt-in).

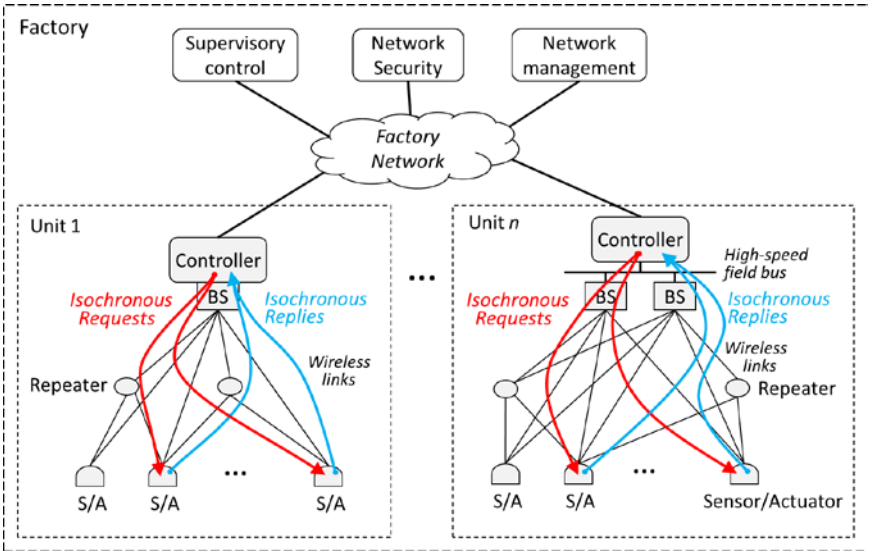


Figure 2.18. Communication Path for Factory Automation.[79]

Process automation is another use case where IIoT is used for continuous production processes to produce or process large quantities or batches of a certain product. Process automation requires deterministic behavior and therefore requires typical latency of 50 milliseconds (ms); can cover relatively large areas meaning wide wireless transmission ranges are required with modest to high connection density; end-to-end throughput, security and availability become more important, but real–time communication requirements decrease; and has smaller numbers of devices, however, each has high throughput requirements for a network with high total capacity. These are requirements that 5G IoT can provide.

A high level of automation requires a higher percentage of 5G IoT connections in industrial automation use cases, as estimated in Figure 2.19 by Ericsson.

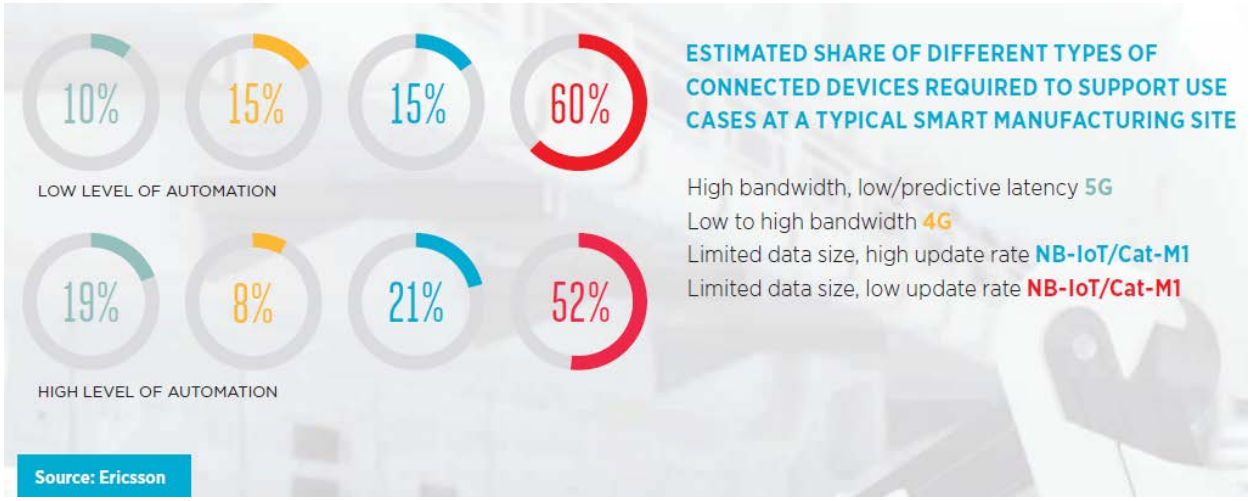


Figure 2.19. Mobile IoT Connection Type Estimation.[80]

[79] *3GPP Technical Specification 22.261 Service requirements for the 5G system Stage 1*, 3GPP. March 2019.
[80] *GSMA Industrial IoT Case Study: Ericsson Smart Factory.* GSMA and Ericsson. September 2018.

By testing the technologies in its own factories, Ericsson found that Mobile IoT and mmWave networks can support a wide range of different manufacturing use cases, making it possible to securely and efficiently optimize manufacturing variables with a single cellular communication system. Cellular networks allow for massive data collection and analytics, increasing intelligent automation on the factory floor and enabling adaptive production. Cellular connectivity also enables fast and cost-efficient production line changes, as well as integration and optimization of contributing workflows.

While, IIoT will initially improve existing processes and augment current infrastructure, the ultimate goal will be to realize entirely new, and dramatically improved products and services. IIoT involves a substantial breadth and depth of technologies, many of which require careful integration and orchestration. According to Research And Markets, leading managed service providers are looking beyond core Machine-to-Machine (M2M) communications towards more advanced services that involve IoT platform and device mediation, data management, and application coordination. As M2M messaging is evolving to a flatter hierarchical structure with edge computing networks, managed privacy and security services will be required to ensure data integrity and asset protection. Data analytics solutions provide the means to process vast amounts of machine-generated, unstructured data captured by M2M systems. As IIoT progresses, there will an increasingly large amount of unstructured machine data. The growing amount of machine generated industrial data will drive substantial opportunities for AI support of unstructured data analytics solutions.[81]

Teleoperation and tele-robotics will leverage the ability to control real machines/equipment by virtual object through master controlling interfaces. Research And Markets sees teleoperation being transformed by digital twin technologies, which refers to the mapping of the physical world to the digital world in which IoT Platforms and Software are leveraged to create a digital representation of physical object or asset. The digital twin of a physical object can provide data about the asset such as its physical state and disposition.

The use of 5G for IIoT networks will be of great importance to certain industry verticals such as agriculture, logistics, and manufacturing. The combination of robotics, teleoperation, and cloud technologies is poised to transform enterprise operations, industrial processes, and consumer services across many industrial related industry verticals. All of these industrial sectors will also require efficient and effective computing systems. There is a substantial opportunity for both a centralized cloud as a service model for software, platforms, and infrastructure as well as edge computing cloud solutions for industry. The combination of robotics, teleoperation, and cloud technologies is poised to transform industrial processes across many industry verticals.

IIoT solutions are already profoundly affecting industrial processes and creating opportunities for product and service transformation. In some cases, entirely new business models will result from the integration of broadband wireless and cloud technologies. In addition, IIoT solutions are evolving from transparency into operations to proactive maintenance and correction.

2.6 SMART CITIES

In smart cities, a massive number of IoT devices will be deployed for several use cases. This could include all applications that gather large amounts of data from sensors and cameras to make the city operation more efficient.

[81] *Global Industrial Internet of Things IIoT Market 2019-2024: Focus on Automotive and Transportation, Cargo and Logistics, Healthcare, Manufacturing, Oil and Gas, & Utilities.* Report by ReasearchAndMarkets. 16 May 2019.

Key challenges to enabling a large-scale uptake of massive IoT include: device costs, battery life, scalability, latency and coverage.[82] Within smart cities and the Internet of Things (IoT), use of 5G technologies, NB-IoT and LTE-M will have a positive impact on these challenges.

In order to ensure long battery lifetime and to reduce energy consumption, these technologies are enabled with two power saving features: extended Discontinuous Reception (eDRX) and Power Saving Mode (PSM). eDRX is a mechanism that enables the device to switch off part of its circuitry to save power. An eDRX cycle consists of an "On Duration" during which the device checks for paging and an "eDRX period" during which the device is in sleep mode. This feature is useful for device-terminated applications, for example, smart grid. PSM is a low-power mode that allows the device to skip the periodic page monitoring cycles, allowing the device to sleep for longer. However, as a result the device becomes unreachable. It is therefore best utilized by device-originated or scheduled applications, for example, smart metering.[83]

The top IoT Smart City applications/projects deployed in 2018 are Smart Traffic Management, and Smart Grids (Utilities) as shown in Figure 2.20, as per IOT Analytics.

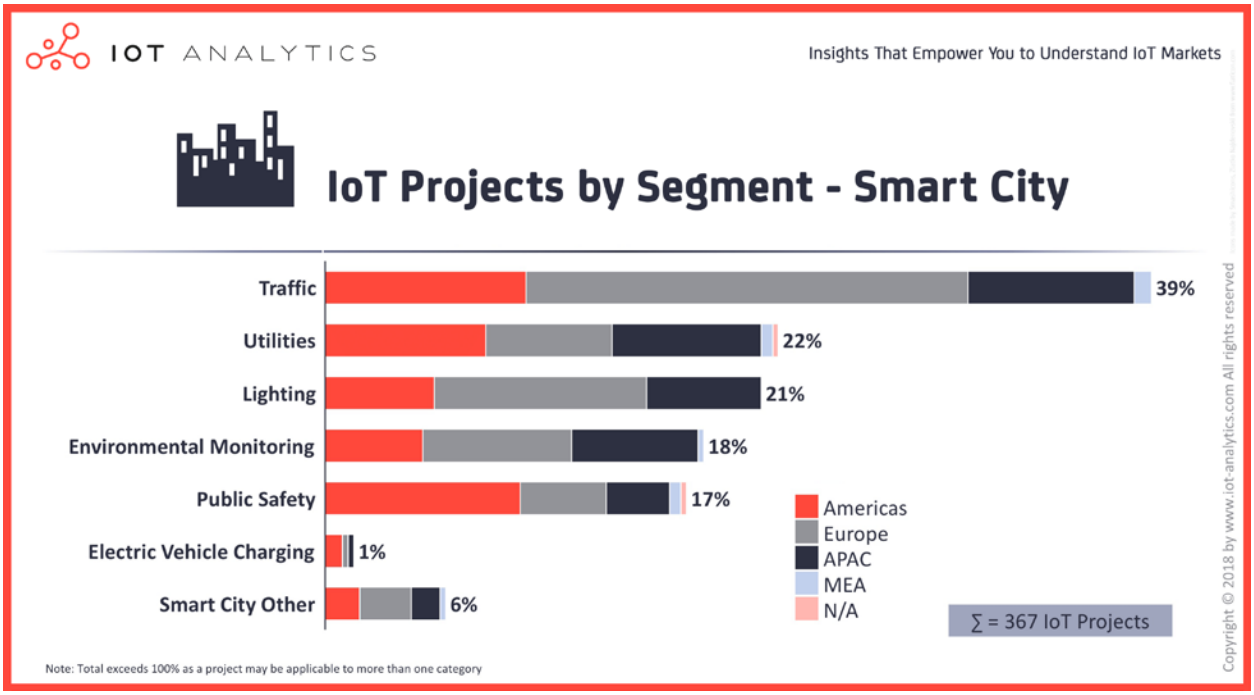


Figure 2.20. IoT Projects by Segment 2018 – Smart City.[84]

2.6.1 SMART TRAFFIC MANAGEMENT

Managing traffic in an effective way is a common application in a Smart Cities project with the support from intelligent street lights equipped with sensor networks to monitor the flow of traffic and help controlling the traffic lights. For example, there could be built-in car navigation systems to guide drivers for nearby places to park and avoid traffic jams. This guidance would decrease traffic congestion, deliver high productivity and improve the quality of life of all the residents.

[82] *Evaluating the Performance of eMTC and NB-IoT for Smart City Applications,* arxiv.org, 20 Nov. 2017.
[83] *Ibid.*
[84] *State of the IoT & Short-Term Outlook 2018*, report by IoT Analytics, 2018.

2.6.2 SMART GRIDS

5G in the smart city concept will also help improve energy systems. 5G will enable very low-cost connections between devices and monitor the devices to properly manage energy needs. The most important part is that it will help in balancing the load of the cities in such a way to reduce electricity peaks, hence reducing the cost of the energy. In the case of power cuts or failures, real time diagnosis will be enabled, for example, shifting the load to a different transformer or device. Another example is making lights dim when there is no public in the vicinity and thereby saving power and reducing energy costs. Smart grids provide a long-term energy saving plan through which energy cost can be reduced.

2.7 ENTERPRISE IOT

The transformation reflected in the IoT is driven by the continuous evolution of technology. Today, everything is digital. Businesses need to evolve with the changing trends and find new and creative ways for enabling success and innovation. However, the ultimate need for enterprises to drive IoT is to enhance operational efficiency, mitigate risks, improve functional visibility, ensure maximum customer engagement, increase revenue streams, and tap into potential opportunities for growth. All these factors can be improved by the real-time information and insights provided by IoT connected things and devices. The 'perfect storm' of increased connectivity (more things), big data and cloud computing (more information), and the automation of core customer and marketing processes is driving enterprise IoT. By automating marketing best practices through IoT, organizations can combine information technology with their current business operations.

5G will play a vital role in enterprise IOT by connecting billion of devices and sharing data independently for a mixed domain of applications/services. Already existing 3GPP and LTE networks remain the most engaging technique for communication in the IOT connectivity, offering IoT systems with a wide area coverage, high security, simplicity in management and access to trusted spectrum. However, the already existing networks cannot support MTC which is essential in IoT. This is where 5G comes into play, providing a solution to the issue. The 5G network can provide the fastest network data rate with relatively low latency and better coverage for MTC communication in relation to present 4G (LTE). The Machine-to-Machine (M2M) communication enables many smart devices and sees a world that is well connected.[85]

The 5G network is capable of supporting massive devices and new services for example, enhanced Mobile Broadband (eMBB), massive Machine Type Communications (mMTC), Critical Communications and Network Operations. 5G mobile system for enterprise IOT trusted to provide low latency, high versatility and high throughput for large number of gadgets, productive energy utilization system and finally connectivity of end clients/devices.[86]

The increasing computing capacity of today's devices allow them to perform complex computations on-site, resulting in edge computing. Edge computing extends cloud computing capabilities by bringing services close to the edge of a network and thus supports a new variety of services and applications.[87]

Security is one of the core requirements of all modern systems. The use of edge computing in IoT is subject to developing secure systems and applications. The security requirement is almost undefined in cloud

[85] *5G Internet of Things: A survey*, L. D. X. S. Z. Shancang Li, Elsevier-Journal of Industrial Information Integration. 2018.
[86] *IoT and 5G: The Interconnection,* Uzairue Stanley, Nsikan Nkordeh, Victor Matthews Olu, Ibinabo Bob-Manuel, Covenant University, Ota, Nigeria. September 2018.
[87] *The role of edge computing in Internet of Things*, Najmul Hassan, Saira Gillani, Ejaz Ahmed, Ibrar Yaqoob, Muhammad Imran, IEEE. May 2018.

computing. Therefore, cloud services are more vulnerable to certain types of attacks. That includes data breaches because of their specific structure. However, in edge computing, security may be well defined and clearly implemented. Therefore, improved data security can be provided because client data are aggregated at certain access points placed close to the end user.

Although enterprise IoT is a relatively new development, in a survey by McKinsey 98 percent of survey respondents reported that most companies within their industry include enterprise IoT initiatives in their strategic road maps as well as improvement in operations, increasing visibility, enabling new business models and creating new product and services as shown in Figure 2.21.[88]

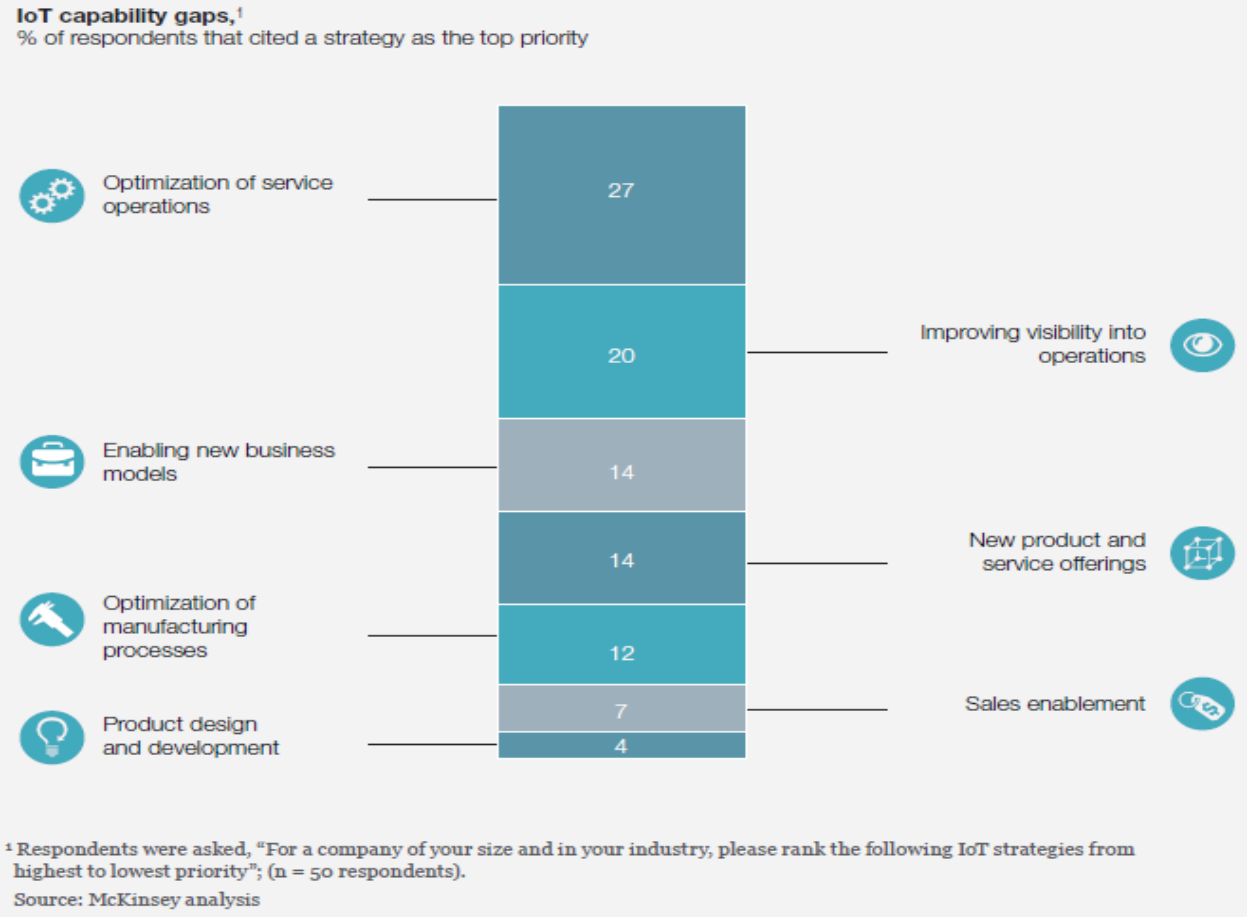


Figure 2.21. IoT Capability Gaps.[89]

[88] *Taking the pulse of enterprise IoT*, McKinsey & Company, High Tech. July 2017.
[89] *Ibid.*

2.7.1 ENTERPRISE IOT BENEFITS AND OPPORTUNITIES

Enterprise IoT can help in improving multiple functions; in a McKinsey survey on where IoT would have the greatest impact, 40 percent of respondents cited improvement in service operations and 30 percent chose manufacturing, as shown in Figure 2.22.[90]

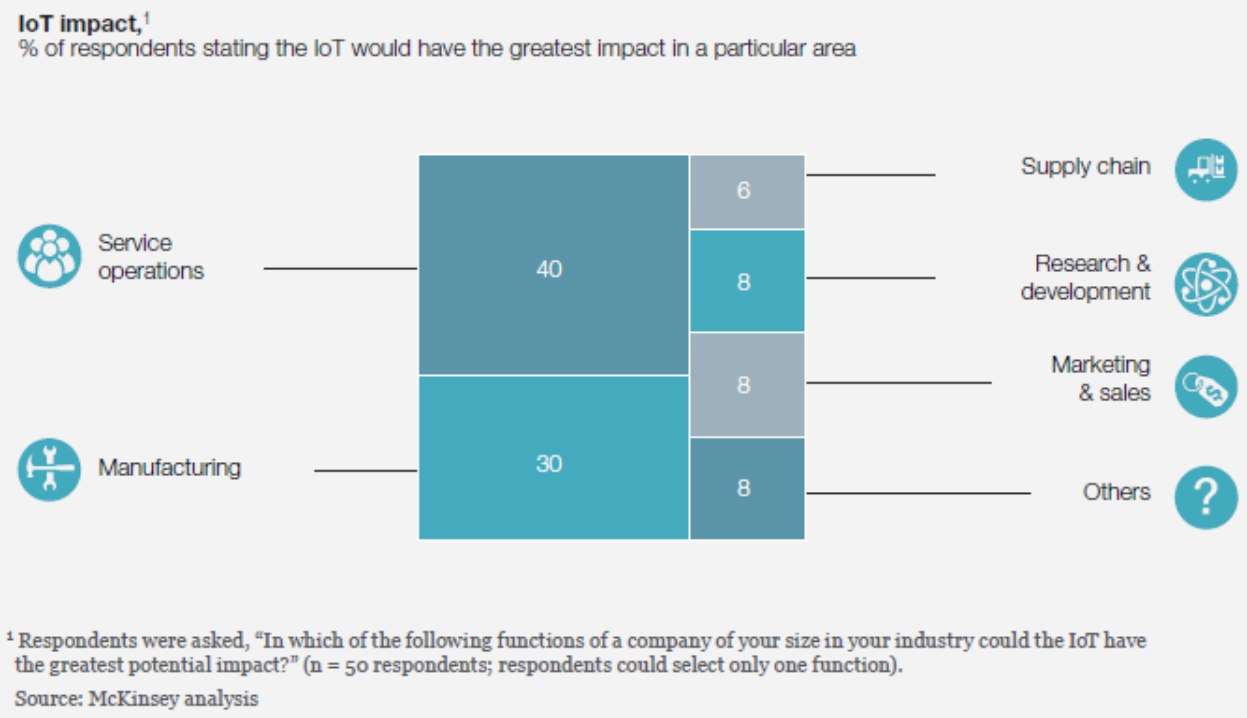


Figure 2.22. Survey Results of Enterprise IoT.[91]

Although enterprise IoT is already providing lots of benefits and is continuously evolving, enterprise IoT capability gaps have been noted in some research. Based on companies already using IoT data, 60 percent surveyed stated that they got significant insights, such as data on customer demographics or shopping patterns. However, an almost equal number ~54 percent claimed that their companies are using only 10 percent or less of this information.[92]

Anecdotal learning from enterprises already using IoT in the 2017 McKinsey study included:

- For one gas rig, managers reported 30 thousand sensors reporting lots of data, and at the time they were only using 1 percent of data for decision making
- 70 percent of the companies had not yet integrated IoT solutions in their existing business workflows
- Boeing workers using IoT wearables and augmented reality tools on wiring harness assembly lines resulted in up to 25 percent improvement in productivity

[90] *Taking the pulse of enterprise IoT*, McKinsey & Company, High Tech. July 2017.
[91] *Ibid.*
[92] *Ibid.*

- Elevator companies were creating a suite of IoT-enabled services to improve the reliability of products and decrease downtime resulting in lower operating costs and transformation of business models

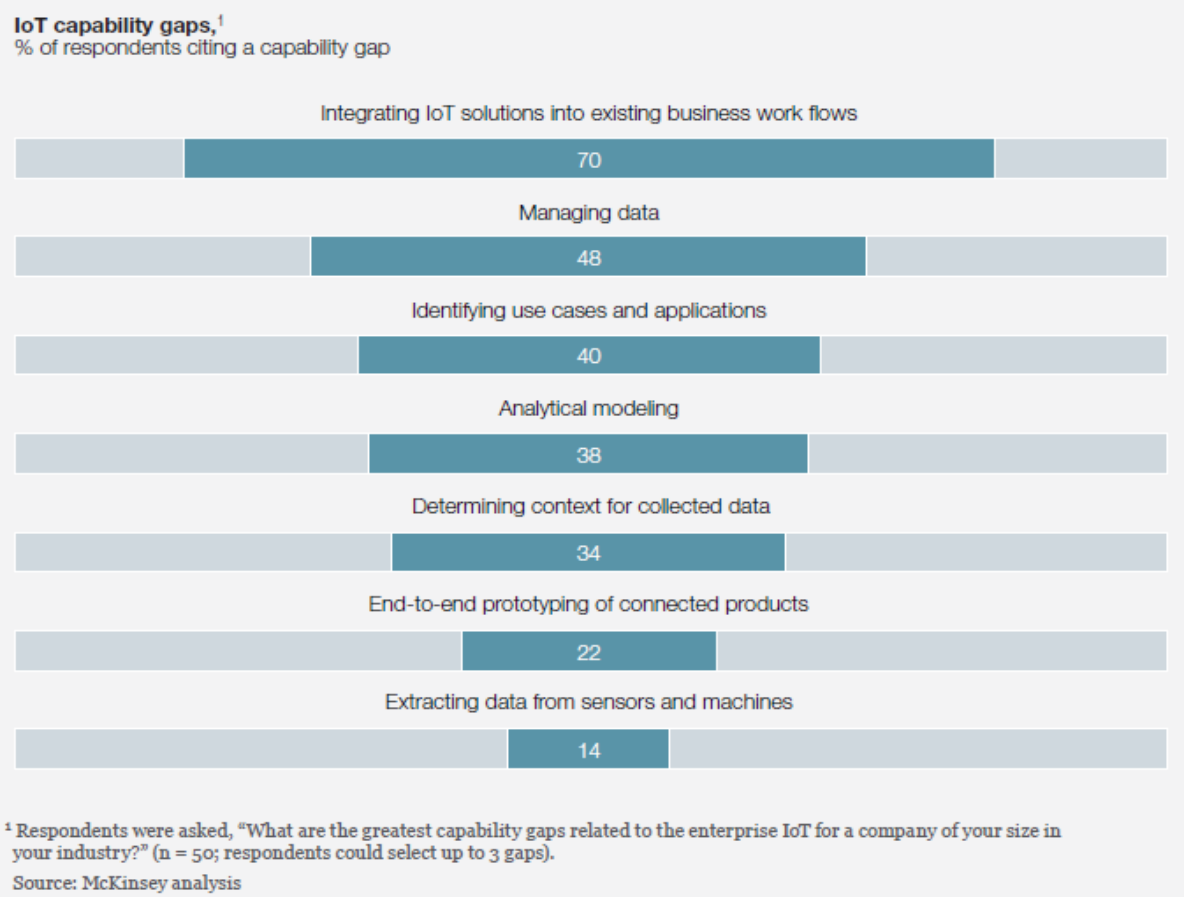


Figure 2.23. Capability Gaps in Companies Using IoT.[93]

The 2017 McKinsey survey pointed to many opportunities for the IIoT to fill the gaps for improved business models across many levels.

2.8 CONSUMER IOT

Consumer Internet of Things applications can range from simple and inexpensive, such as personal fitness devices to high-end smart home automation applications. Thus, IoT use cases, devices and applications for consumers are very diverse as well and sometimes it is difficult to differentiate between consumer and industrial/enterprise IoT. For example, wearables are typically considered as personal use and therefore consumer IoT devices, however, some wearables are developed specifically for use in factories or mines for worker safety with expanded durability. The point being, just as it may be difficult to draw a clear line between Industrial and Enterprise IoT, the line for Consumer IoT may also be blurred. Research firms have a variety of ways to categorize the IoT. Several examples are provided in this report.

[93] *Taking the pulse of enterprise IoT*, McKinsey & Company, High Tech. July 2017.

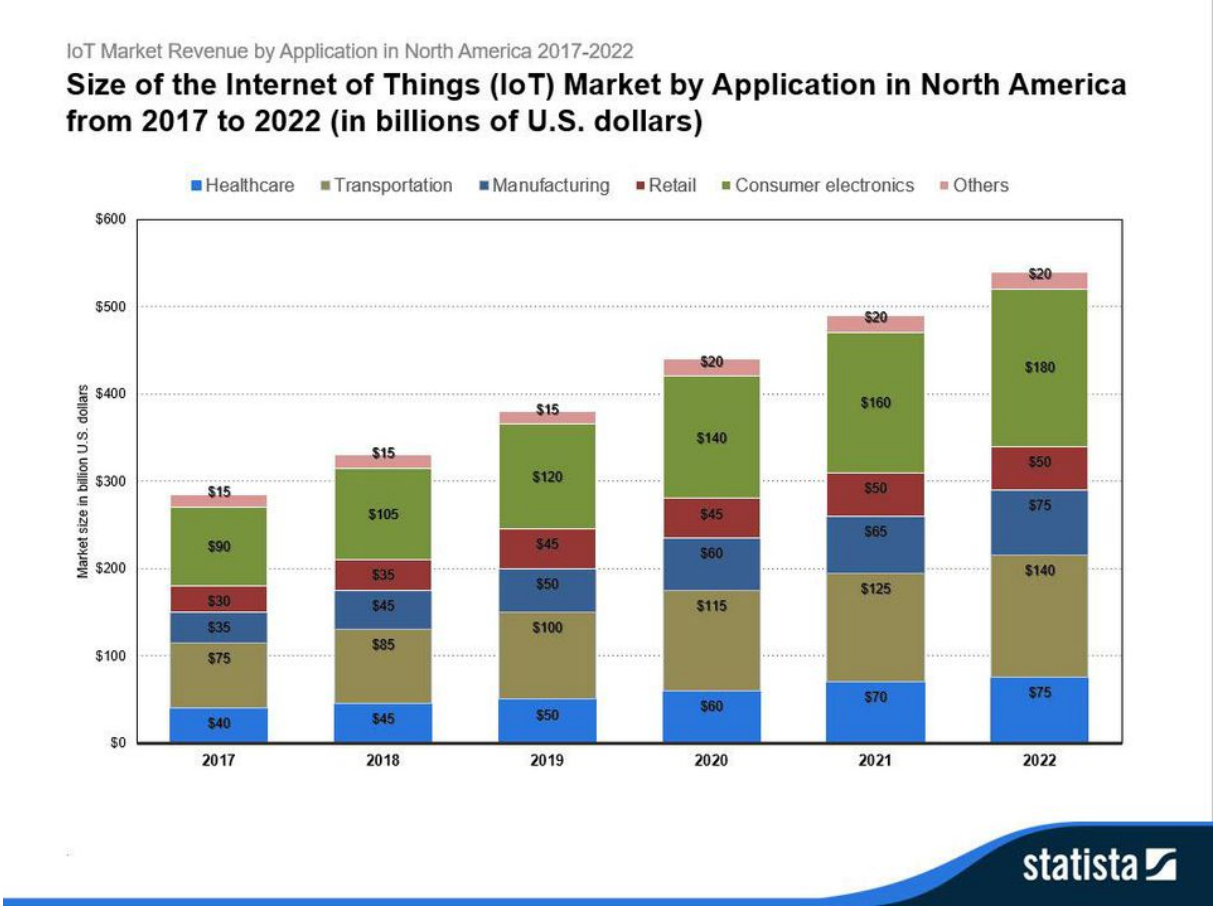


Figure 2.24. IoT Market by Application in North America: 2017-2022 (Billions).[94]

Figure 2.24 from Statista defines the North America IoT market by application. Consumer electronics would entail the consumer IoT segment. The North American IoT consumer electronics market is predicted to increase from $90B in 2017 to $180B in 2022, attaining a CAGR of 12.25 percent.[95]

Some of the more popular consumer IoT devices and applications, according to end-of-2016 research conducted by the Interactive Advertising Bureau (IAB),[96] indicated the highest consumer buying interest in areas such as: connected and smart TV, and connected streaming devices; connected car applications; wearables; home control devices and systems; voice-command systems; IoT-enabled appliances; virtual reality (headsets and games); and all types of smart watches. Buying interest in smart glasses was still relatively small (see Figure 2.25).[97]

[94] *Ten Charts that will Challenge your Perspective on IoT Growth*. Forbes. 6 June 2018. Source: Statista.
[95] *Ibid.*
[96] The Interactive Advertising Bureau (IAB) empowers the media and marketing industries to thrive in the digital economy. Its membership is comprised of more than 650 leading media and technology companies that are responsible for selling, delivering, and optimizing digital advertising or marketing campaigns. 11 February 2019.
[97] *Consumer Internet of Things (CIoT) – what is it and how does it evolve?* I-Scoop. Source: IAB.

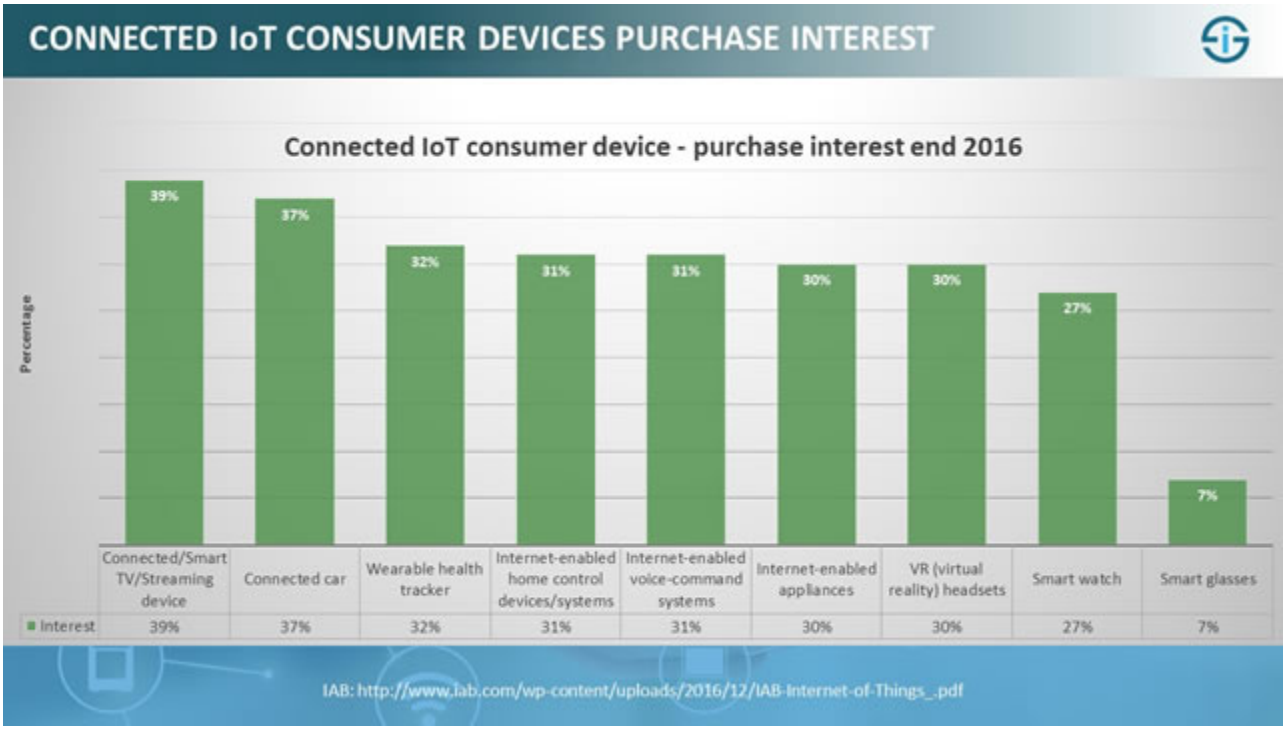


Figure 2.25. Connected IoT Consumer Device Buying Interest.[98]

IDC Research Director Marcus Torchia stated that Consumer IoT will become the second largest industry segment in terms of IoT spending globally in 2019 by reaching $108 billion. The leading consumer use cases will be related to the smart home (home automation and smart appliances in particular), personal wellness, and connected vehicle infotainment. Consumer IoT will be the fastest growing industry segment overall with a five-year CAGR of 17.8 percent.[99]

It is forecast by Statista that 12.86 billion IoT sensors and devices will be in use in the consumer segment by 2020, growing at a 35 percent CAGR per year from 2017. Vertical-specific sensors and devices are projected to grow from 1.64 billion units in 2017 to 3.17 billion in 2020, attaining a 24.57 percent CAGR in just three years.[100]

2.8.1 SMART HOMES

Smart home, also called home automation, is an automation system that controls the lighting, climate, entertainment, appliances, and home security such as alarm systems of a household. The home automation market is exploding with the need for personal security, energy conservation and information/entertainment with everything from the smart TV, to sophisticated HVAC systems, to the video surveillance doorbell. Home automation systems will make use of the 5G high speed connection between devices. Daily home usage devices, including air conditioners, refrigerators, lighting, heaters, washers, dryers and etcetera can be

[98] *Consumer Internet of Things (CIoT) – what is it and how does it evolve?* I-Scoop. Source: IAB.
[99] *IDC forecasts worldwide spending on the Internet of Things to Reach $745 billion in 2019,* led by the manufacturing, consumer, transportation, and utilities sectors, press release by IDC, 3 January 2019.
[100] *IoT Installed Based by Category 2014-2020.* Statista. https://www.statista.com/statistics/370350/internet-of-things-installed-base-by-category/.

connected with each other and controlled remotely using a smart phone thereby saving both time and money. Surveillance cameras could be very effective for home security purposes and improve the quality of life of all the residents.

The overall global smart home market is forecast to grow from USD 76.6 billion in 2018 to USD 151.4 billion by 2024, at a CAGR of 12 percent.[101] According to MarketsAndMarkets and other research firms, the growth of the smart home market is driven by various factors:

- large base of internet users and increased adoption of smart devices
- rise in the awareness of fitness and healthy lifestyles owning to the high disposable income in developing economies
- high importance of home monitoring from remote locations
- rise in the need for energy- saving and low carbon emission solutions
- cost reduction measures enabled by smart homes
- rapid proliferation of smartphones and smart gadgets
- existence of market players focusing on expanding their smart home product portfolios
- wide-spread concern about safety, security, and convenience

The global market for smart home devices is expected to grow about 27 percent year-over-year in 2019 to nearly 833 million shipments, according to IDC. Sustained growth is expected to continue with a compound annual growth rate (CAGR) almost 17 percent over the 2019-2023 forecast period and nearly 1.6 billion devices shipped in 2023 as consumers adopt multiple devices within their homes and as global availability of products and services increases.[102]

Table. 2.3. Smart Home Devices by Category, 2019 and 2023 (shipments in millions).[103]

Product Category	2019 Shipments*	2019 Market Share*	2023 Shipments*	2023 Market Share*	2019 – 2023* CAGR
Home Monitoring/Security	140.3	16.8 percent	351.7	22.6 percent	25.8 percent
Lighting	56.9	6.8 percent	183.2	11.8 percent	34.0 percent
Others	114.3	13.7 percent	269.4	17.3 percent	23.9 percent
Smart Speaker	144.3	17.3 percent	240.1	15.4 percent	13.6 percent
Thermostat	18.8	2.3 percent	37.5	2.4 percent	18.8 percent
Video Entertainment	358.1	43.0 percent	475.4	30.5 percent	7.3 percent
Total	**832.7**	**100.0 percent**	**1,557.4**	**100.0 percent**	**16.9 percent**

*** Note**: All data represents forecast values

Smart home appliances are sometimes the only option for consumers to buy today, similar to smart TVs. This drives appliances as a key sector in the smart home environment. All remote-control appliances used to have very limited range meaning the user needed to have proximity to operate with the remote control. However, that is no longer applicable; linked devices can be operated from almost anywhere in the world in a real time scenario. This has value in terms of flexibility, safety, convenience and energy savings. Research by Clutch compares three popular devices that people use the most in Figure 2.26.

[101] Smart Home Market by Product (Lighting Control, Security and Access Control, HVAC, Entertainment, Smart Speaker, Home Healthcare, Smart Kitchen, Home Appliances, and Smart Furniture), Software & Services, and Region – Global Forecast to 2024, MarketsAndMarkets. 2019.
[102] IDC Worldwide Quarterly Smart Home Device Tracker, IDC. 29 March 2019.
[103] *Ibid.*

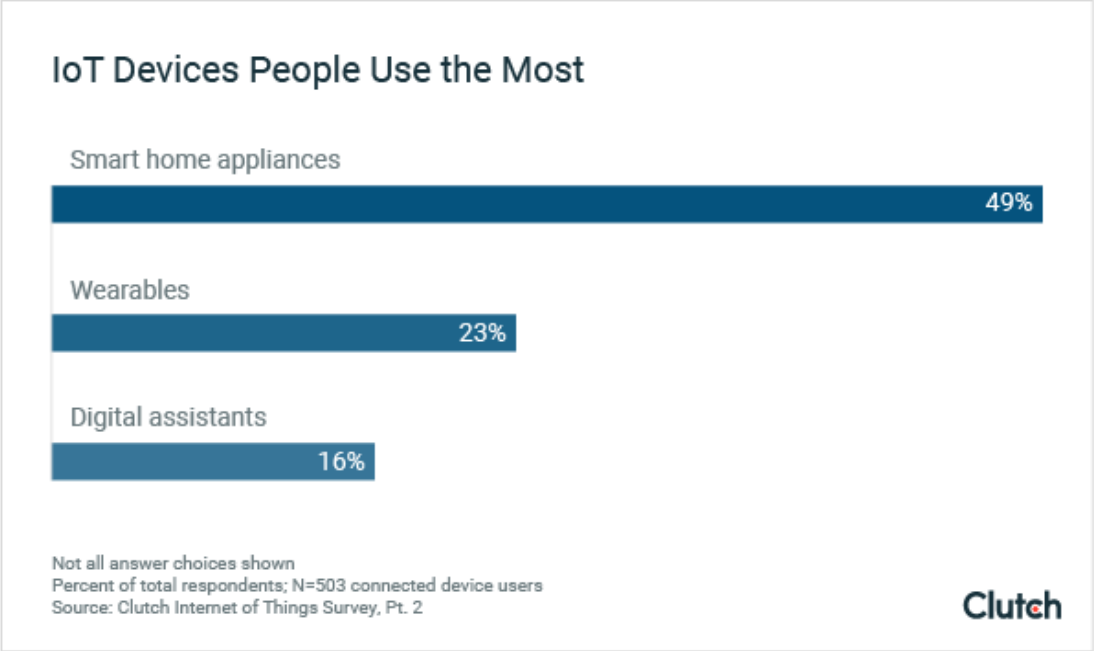


Figure 2.26. IoT Devices People Use the Most.[104]

2.8.2 WEARABLES

Wearables are smart electronic devices (with micro-controllers) that can be worn or incorporated into clothing such as activity trackers/fitness bands, eyeglasses and smart watches. Typically, the wearable will enable objects to exchange data through the internet with a manufacturer, operator or other connected device without human intervention.[105]

Personal healthcare is an important industry application for wearables with devices to support blood pressure, heart rhythm, glucose monitoring, temperature and other body measurements with IoT connections directly to physicians or other healthcare professionals. Alerts can be provided and there are already use cases where timely notifications have saved peoples' lives. Other areas of use in the healthcare industry include patient surveillance, fall detection and etcetera. Wearable technology has a bright future in healthcare with the passive monitoring of vital statistics.

Monitoring of vital signs is another use case for the fitness trackers that are extremely popular for everything from measuring footsteps to heartrate to speed.

Other devices that may fit into wearables are tracking devices, such as trackers inserted into children's backpacks or clothing, trackers on pet collars or trackers for any type of personal equipment or possessions. Asset tracking is enabled over long distances where IoT coverage is present for LPWA networks.

Headsets that can be used by consumers or in industrial or enterprise settings would also fit into the wearables category.

[104] *Do Wearable Devices Connect People to the Internet of Things?* Article by Grayson Kemper, Clutch. 15 November 2018.
[105] Wikipedia.

Connected wearables grew at a good pace in Q1 2019 continuing the surprisingly good run in 2018 according to Chetan Sharma Consulting.[106]

2.8.3 CONNECTED CAR

A connected car is a car that is equipped with Internet access, and usually also with a wireless local area network. This allows the car to share internet access, and hence data, with other devices both inside as well as outside the vehicle.[107] Connected cars will be enabled by 5G networks through Vehicle-to-Everything or V2X technology, enabling communication between cars, buses, trucks, trains, roads, pedestrians and road infrastructure, and our smartphones and other devices.

The U.S. Department of Transportation's (USDOT's) Connected Vehicle program is working with state and local transportation agencies, vehicle and device makers, and the public to test and evaluate technology that will enable vehicles to "talk" to everything. Connected vehicles could dramatically reduce the number of fatalities and serious injuries caused by accidents on our roads and highways.

A National Highway Traffic Safety Administration (NHTSA) study of connected vehicle technologies has shown the potential to reduce up to 80 percent of crashes (where drivers are not impaired), which would save a significant number of lives and prevent millions of crash-related injuries every year. While the number of people surviving crashes has increased significantly thanks to airbags, anti-lock brakes, and other technology, the USDOT is shifting its focus from minimizing injury from crashes to prevention of accidents. Every year, there are more than 5 million accidents and over 30,000 fatalities with many additional serious injuries, according to the NHTSA. In fact, car accidents are the leading cause of death among young children and young adults according to the Centers for Disease Control.[108]

V2X will offer advantages over technologies now appearing in high-end vehicles, such as radar, lidar, cameras, and other sensors. For one thing, V2X will offer greater range than on-board vehicle equipment, allowing alerts of hazardous situations much earlier, providing more time to react and prevent an accident.

In addition to the tremendous safety potential of connected vehicles, they also promise to increase transportation options and reduce travel times. Traffic managers will be able to control the flow of traffic more easily with the advanced communications data available and prevent or lessen developing congestion. This could have a significant impact on the environment by helping to cut fuel consumption and reduce emissions.

Connected cars are already contributing significantly to the bottom line of service providers. For example, at the 4th quarter of 2018, AT&T added more than 1.5 million cars to its network for the ninth straight quarter for a total of 32 million connected vehicles.[109]

5G Americas white paper, *Cellular V2X Communications Towards 5G* describes the use cases, starting with the advanced driving categories identified in 3GPP, including ranging/positioning, extended sensors, platooning and remote driving. The paper also describes how mobile network operators, vehicle

[106] *US Mobile Market Update – Q1 2019*, report by Chetan Sharma, Chetan Sharma Consulting. May 2019.
[107] Definition by Wikipedia.
[108] *Connect Vehicle Basics*, U.S. Department of Transportation website. May 2019.
https://www.its.dot.gov/cv_basics/cv_basics_what.htm
[109] *US Mobile Market Update – Q1 2019*, report by Chetan Sharma, Chetan Sharma Consulting. May 2019.

manufacturers, cloud service providers and regulatory bodies can work together to deliver a brand-new experience for drivers, travelers and other road users in the near future.

V2X was introduced with 802.11p and supported a limited set of basic safety services. With 3GPP Release 14, V2X could expand to support a much wider, richer range of services: from low-bandwidth safety applications to high-bandwidth applications such as passenger infotainment. 3GPP Releases 15 and 16 will enable even more V2X services by providing longer range, higher density, very high throughput and reliability, highly precise positioning and ultra-low latency with 5G technology. Figure 2.24 summarizes these features and shows how 802.11p, LTE and 5G may coexist for some time, depending on the region. It is important to note that the 5G radio access enhancements will enable advanced use cases for data exchange but will not duplicate the 4G-based V2X functionality. This way, 5G V2X services are additive to the foundational capabilities of LTE V2X. Indeed, 5G will be future-proof and backwards compatible with LTE V2X. A more detailed explanation of the requirements for V2X communication in Release 17 standards is described in section 3.1.1.14 of this whitepaper.

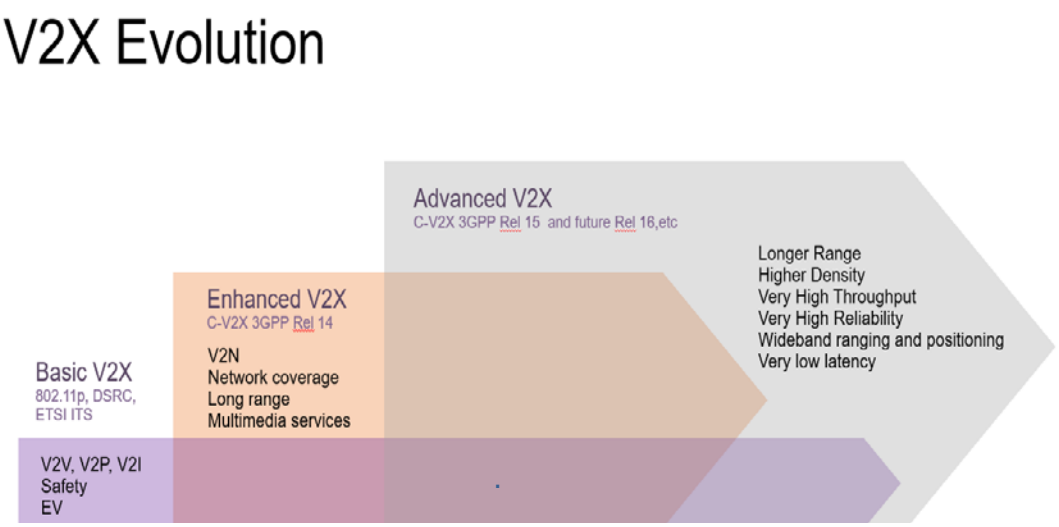


Figure 2.27. V2X Evolution.[110]

IDC estimates that worldwide shipments of connected vehicles, which includes options for embedded and aftermarket cellular connectivity, will reach 51.1 million units in 2019, an increase of 45.4 percent over 2018. By 2023, IDC expects worldwide shipments to reach 76.3 million units with a five-year CAGR of 16.8 percent.[111] The sustained growth of the connected vehicles market is being driven by a multitude of factors, including consumer demand for a more immersive vehicle experience, the ability of auto manufacturers to better utilize connected vehicles for cost avoidance and revenue generation, evolving government regulations, and mobile network operator investments in new connections and services. "The automotive ecosystem is positioning the vehicle as the next, emerging digital platform," said Matt Arcaro, research manager, Next-Generation Automotive Research at IDC. "Deploying embedded or aftermarket connectivity at scale will be key to unlocking its potential."[112]

[110] *Cellular V2X Communications Towards 5G*, white paper by 5G Americas. March 2018.
[111] *Worldwide Connected Vehicle Shipments Forecast to Reach 76 Million Units by 2023, According to IDC,* press release. 23 May 2019.
[112] *Ibid.*

2.9 IOT MARKET OVERVIEW - CONCLUSION

Globally, more than half of the things that will be connected to IP networks will not be our personal devices (smartphone, tablets, PCs and TVs), but sensors, tracking modules, cameras and other forms of machine-to-machine (M2M) connections that gather and share various types and volumes of information with other machines. Up until 2018, the traffic generated from IoT applications was practically negligible (less than 5 percent of global traffic). With more bandwidth-intensive (and low latency) applications like autonomous driving smart cars, video surveillance, and connected health, IoT traffic now needs to be accounted for, secured and managed in new ways. In addition, there is new and emerging monetary value that will be generated by IOT applications potentially creating much-needed revenue streams that may further support IoT and network innovation.

3. IOT REQUIREMENTS

Compared to traditional human type communications, the Internet of Things have many different requirements. First, the IoT devices should have low complexity and low cost due to the large scale of deployments. Second, they should have very long battery life to avoid the hassle of changing or recharging millions of IoT devices frequently. Depending on the type of IoT services, some IoT devices might be installed in areas with poor cell coverage, such as basements or deeply inside a storage building, extended coverage will be needed to ensure reliable services. Last but not least, as more and more new IoT applications continue to emerge, the number of IoT connectivity links are expected to grow exponentially. The IoT network should be scalable to support a large number of simultaneously connected IoT devices without sacrificing the network performance for both human users and machines. To legacy operators, IoT services should ideally to be able to leverage their existing cellular infrastructure and co-exist with other non-IoT services in the same network and spectrum.

3.1 3GPP REQUIREMENTS

Support for Massive IoT has been a key consideration of 3GPP's approach to 5G standards. Building on the base of LTE-based NB-IOT and eMTC standards, 5G expands the capacity and capabilities to serve the growing use of IoT devices in industry, enterprise, and the home. Significant specific enhancements in both the 5G core network and radio are underway in Rel-16 to support both the increasing number of IoT devices and to meet the increasingly stringent diverse requirements coming from new markets. Continuing emphasis on support of the IoT market is seen in Rel-17 and beyond.

3.1.1 RELEASE 16

The 5G network is expected to be able to provide optimized support for a variety of different services, different traffic loads and different end user communities. The system should be able to simultaneously support multiple services with different requirements of reliability, latency, throughput, positioning and availability. Some new IoT services, such as Unmanned Aerial Vehicle (UAV) control and factory automation, have very stringent requirements on latency, reliability and positioning. In some other massive IoT use cases, the 5G system is expected to support tens of millions of devices sending and receiving data simultaneously, which will require enhanced connection modes and evolved security mechanism.

In order to have flexible network operations, the 5G network will need to support network slicing, network capability exposure, scalability and diverse mobility. To meet the stringent requirements on latency and reliability, the 5G network optimizes the resource efficiency of both control and user plane and minimizes

routing between end uses and application servers. Enhanced charging and security mechanisms will be needed to handle new types of IoT devices that are connected to the network in different ways. In addition, advanced IoT services like emerging V2X applications also have much more stringent requirements on data rate, reliability, latency, communication range, and speed. 3GPP has completed Release 16 service requirements for different services, such as eMBB, IoT and URLLC. This section summarizes the relevant requirements for 5G IoT provided by 3GPP Release 16 standards. For detailed requirements, please refer to:

- 3GPP TS 22.261 V16.7.0, *Service requirements for the 5G system*; Stage 1. March 2019
- 3GPP TS 22.186 V16.1.0, *Enhancement of 3GPP support for V2X scenarios; Stage 1*. December 2018
- 3GPP TS 22.104 V16.1.0, *Service requirements for cyber-physical control applications in vertical domains; Stage 1*. March 2019

The following sub-clauses describe the Rel-16 requirements that have been included in 3GPP TS 22.261. As the downstream working groups of 3GPP are completing their work on Rel-16, it is apparent that not all of these will be met by the current stage 2/3 efforts. Requirements that are not met in Rel-16 stages 2 and 3 will be carried forward to Rel-17.

3.1.1.1 SUPPORT FOR NETWORK SLICING

Network slicing allows the operator to provide flexible services based on different customer requirements on functionality, performance or specific group of users. Each network slice can provide the functionalities of various parts of the network, including radio access network, core network and IP Multimedia Subsystem (IMS) functions. One or several network slices can be supported at the same time.

The 5G network allows the operator to create, modify and delete a network slice. The operator can define and update the set of services and capabilities supported in a network slice. The network supports scaling of a network slice and adaptation of its capacity based on needs. A priority order between different network slices can be defined in the event that multiple network slices compete for resources on the same network. An unauthorized UE is prevented from accessing a radio resource dedicated to a specific private slice for any purpose other than that authorized by the associated third party. Cross-network slice coordination is supported so that the selected services of a non-pubic network can be extended through a Public Land Mobile Network (PLMN).

3.1.1.2 SUPPORT FOR DIVERSE MOBILITY MANAGEMENT

A key feature of 5G is to support UEs with different mobility management needs. The 5G network supports IoT devices with a wide range of mobility management needs, including:

- Stationary during their entire usable life (for example, sensors embedded in infrastructure)
- Stationary during active periods, but nomadic between activations (for example, eHealth)
- Mobile within a constrained and well-defined space (for example, in a factory)
- Fully mobile (for example, V2X)

Different mobility management methods are supported based on different mobility patterns to minimize signaling overhead and optimize access for different types of IoT devices. Inter- and/or intra- access technology mobility procedures are supported with minimum impact to the user experience.

3.1.1.3 SUPPORT FOR MULTIPLE ACCESS TECHNOLOGIES

The 5G system supports multiple 3GPP access technologies, including New Radio (NR), LTE as well as non-3GPP access technologies such as satellite access, Wireless Local Area Network (WLAN) access and fixed broadband access. For optimization and resource efficiency, the 5G network will select the most appropriate access technologies for an IoT service, potentially allowing multiple access technologies (for example, NR, LTE, non-3GPP) to be used simultaneously for one or more IoT services active on an IoT device.

Based on operator policy, the 5G system can steer an IoT device to select certain 3GPP access technology (for example, NR or LTE). An IoT device can select, manage and efficiently provision services over different access technologies. If a device is using two or more access technologies simultaneously, the 5G network can distribute traffic load optimally between the access technologies. Operation in both licensed and unlicensed bands is supported. Seamless handover between NR and LTE is supported.

3.1.1.4 SUPPORT FOR OPTIMIZATION OF RESOURCE UTILIZATION EFFICIENCY BASED ON DIFFERENT IOT USE CASES

The wide range of 5G IoT applications supported by the 5G system have totally different services and Key Performance Indicator (KPI) requirements. Some applications will require high data rates and very low latency, while some other use cases will require simultaneous connection for a massive number of devices. The 5G network is designed in a way that the system can be optimized flexibly based on individual use cases to offer improvement in efficient resource utilization. Depending on the potential service and network operation requirements, the efficient configuration, deployment and use of IoT devices can be achieved through the following strategies:

- Bulk provisioning to support a massive number of IoT devices more efficiently
- Improving control plane resource efficiencies and reducing signaling overhead based on use case and traffic pattern
- Optimizing user plane resource utilization to achieve low latency and higher reliability
- Optimization for User Equipment (UE) originated data transfer
- Improving efficiencies based on the reduced needs related to mobility management for stationary devices or devices with restricted range of movement

The 5G system optimizes control and user plane resource utilization based on device communication pattern (send-only, frequent or infrequent), mobility pattern (stationary or full mobility), characteristics of payload (small or large payload size), density of connections (massive or limited number of simultaneous IoT devices), timing pattern of data transfer (real time or non-real time). Efficient bulk operation and efficient management of IoT devices is supported for the high density of IoT connections.

3.1.1.5 SUPPORT FOR EFFICIENT USER PLANE

To support a wide range of IoT use cases with different performance requirements and data traffic models such as Internet Protocol (IP) or non-IP, or short bursts or data transmission with high throughput, the 5G network needs to have an efficient user plane. A Service Hosting Environment can be set up inside the operator's network to offer Hosted Services closer to the end user and meet the low latency requirement. User plane paths can also be selected or changed to improve the user experience when User Equipment (UE) or an application changes location during active communication. The 5G network maintains the user experience (for example, Quality of Service (QoS) or Quality of Experience (QoE)) when a UE in an active

communication moves from a location served by a Service Hosting Environment to another location served either by a different Service Hosting Environment or by an application server located outside the operator's network.

3.1.1.6 SUPPORT FOR PRIORITY, QOS AND DYNAMIC POLICY CONTROL

In addition to regular commercial IoT services, the 5G network will also support regional or national regulatory services that might require special priority treatment. Some of these services might share the same QoS requirements (such as latency and packet loss rate) but have different priority requirements. For example, UAV (Unmanned Aerial Vehicle) and air traffic control might have the same stringent requirements on latency and reliability but have different priorities. It is therefore important for the 5G network to support a flexible mechanism that can decouple the priority of a communication from the associated QoS requirement so that services with the same QoS characteristics can be supported with different priorities.

The 5G system can make or change priority decisions a based on the network conditions (for example, during disaster events or network congestion). Required QoS (for example, reliability, latency and bandwidth) is provided for a service, and resources may be prioritized when needed. The network allows multiple services to co-exist, including multiple top priority services. Meanwhile, the network also prevents a single service from consuming or occupying all available network resources, which might impact the QoS of other services competing for resources on the same network. End-to end QoS control (including RAN, core, backhaul and network to network interconnect) is needed to deliver the 5G user experiences. In the case of networks with multiple access technologies, the 5G network can select the right combination of access technologies to serve the UE based on the targeted priority, pre-emption, QoS parameters and access technology availability.

In addition, the 5G network can create and enforce the prioritization policy for users and traffic either during the connection setup or once connected. Optimized signaling for the prioritized users and traffic is supported. Based on operator policy, the 5G network allows real-time, dynamic, secure, and limited interaction with the QoS and policy framework for modification of QoS and policy frame work by authorized users.

3.1.1.7 SUPPORT FOR MULTIPLE CONNECTIVITY MODELS

Depending on the IoT service types, the UE can connect to the 5G network directly, or connect using another UE as a relay UE, or connect using both types of connections. The 5G network can support both direct and indirect connection modes simultaneously. The UEs can be wearable devices monitoring biometrics, or non-wearable devices that communicate in a private LAN such as smart home appliances. The relay UE can get access to the network using 3GPP or non-3GPP access technologies (such as WLAN or fixed broadband access). The 3GPP Radio Access Technology (RAT) can be supported either over licensed or unlicensed bands. When a remote UE is attempting to establish an indirect connection, there might be several relay UEs available for connection. In this case, the remote UE can discover and select the appropriate relay UE and establish the connection between the remote and relay UE. Meanwhile, the 5G network supports the QoS for a user traffic session between the remote UE and the network using 3GPP access technology.

3.1.1.8 SUPPORT FOR EXPOSURE OF NETWORK CAPABILITY TO A TRUSTED THIRD PARTY

In many 5G-IoT use cases, a third party is involved when providing the services. It is necessary to allow the operator to expose some network capabilities to a trusted third party so that the third party can customize a dedicated physical or virtual network or dedicated network slice for a specific use case; and manage a trusted application in a Service Hosting environment. For this reason, the 5G network provides suitable APIs to allow a trusted third party to create, modify or delete network slices used for the third party. Based on operator's policy, a trusted third party can define and update the set of services and capabilities supported in a network slice used by the third party. The 5G network also provides suitable Application Protocol Interfaces (API) to allow a trusted third party to monitor the status (for example, locations, lifecycle, registration status) of its own devices. The number of the supported devices can range from a few to several tens of thousands. The third-party infrastructure may be used in a private slice and the third party can manage its own non-public network and private slices in the PLMN in a combined manner.

3.1.1.9 SUPPORT FOR EFFICENT NETWORK RESOURCE UTILIZATION AND NETWORK OPTIMIZATION BASED SYSTEM INFORMATION

A variety of sensors can be integrated into IoT devices. Different IoT applications running on the devices can also have different communication patterns. In addition, IoT devices might use different access technologies for different applications. The 5G network supports efficient network resource utilization and network optimization based on the system information gathered by sensors, including

- Network conditions, such as network load and congestion information
- Information on serviced IoT devices, such as access technologies (3GPP or non-3GPP); cell type (macro or small cell), user experienced data rate
- Application's communication pattern (expected traffic over time)
- Information on prioritized communication (priority level, traffic associated with the service)
- Mobility information of the devices (stationary, restricted mobility or full mobility)
- Location information of the devices
- Sensor-level information (UE direction, speed, power, etcetera)
- Application-level information (foreground or background apps, user setting)

3.1.1.10 SUPPORT FOR FLEXIBLE BROADCAST/MULTICAST SERVICE

Broadcast/multicast IoT applications in massive IoT use cases will require a flexible and dynamic radio resource allocation between unicast and multicast services within the 5G network or a stand-alone network deployment for broadcast/multicast services. Support for zero to 100 percent of the radio resources of one or more radio carriers will need to be reserved for the delivery of broadcast/multicast content. The operators can configure and broadcast multiple quality levels of the broadcast/multicast content for the same service in a stand-alone 3GPP based broadcast/multicast system. In addition, the 5G network will support parallel transfer of multiple IoT services to a device, or parallel transfer of multiple quality levels of the content to the same UE. The 5G system can also support a stand-alone multicast/broadcast network that consists of multiple cells with inter-site distances of up to 200 kilometers (km). Even though not much work has been done to support flexible broadcast/multi-cast service in Release 16, this subsection provides the requirements of future 5G-IoT systems as broadcast/multi-cast service will be needed to support 5G IoT in the future. MBMS support in 5G is currently being discussed in SA and RAN working groups as part of Release 17.

3.1.1.11 SUPPORT FOR DYNAMIC SUBSCRIPTION WITH THE 5G NETWORK

As the market of the IoT grows exponentially, it is expected that there will be a wide range of different types of IoT devices with different usage models. Different subscription status will be required. For example, an IoT device can be added to an existing subscription or become part of a new subscription; it can also be leased, and its ownership can be changed several times over its life cycle. It is therefore necessary for the 5G network to support a dynamic subscription management mechanism so that the subscription can be modified when the ownership of an IoT device changes. In addition, the 5G system should be able to update or refresh subscriptions when there is a security threat on the credentials of the devices.

As mentioned previously, 5G IoT devices can be connected to the network directly or connected to the network indirectly through a relay UE which can access the network using either 3GPP or non-3GPP access technologies. To identify and manage the IoT devices, the 5G system provides a subscription with the network whether the access is done via 3GPP or non-3GPP access. An IoT device connected to the network directly must have a 3GPP subscription. If an IoT device is connected to the network indirectly through a relay UE and needs to be identified by the core network, it will also be provided with a 3GPP subscription. In addition, depending on operator policy, the 5G system supports a mechanism to provision on-demand connectivity based on needs so that a user can provide its identification securely and be served spontaneously on the spot. The 5G system also supports a secure mechanism for a home operator to provision the 3GPP credentials of an IoT device remotely.

3.1.1.12 SUPPORT FOR OPTIMIZATION OF ENERGY EFFICIENCY

Energy efficiency is a very critical issue for 5G IoT. For many use cases, IoT devices are installed in remote sites, and there are tens and thousands of them. Service maintenance becomes impossible if the batteries of millions of IoT devices have to be changed or recharged frequently. In addition, devices with a small form factor usually have a small battery, which also limits the battery life. 5G IoT devices should have more than a 10-year battery life to ensure long term reliable service. To achieve this goal, the 5G system meets the following requirements:

- Supports Power Saving Mode (PSM). PSM can be activated / deactivated either manually or automatically. When in PSM mode, the devices transmit power may be reduced or turned off (deep sleep mode) as long as the potential increased latency or jitter do not impact the applications
- Supports mechanisms in the system to save battery life of either IoT device or relay device. For example, mechanisms to reduce signaling transmission or reduce paging procedure to reduce device power consumption.

3.1.1.13 SUPPORT FOR MULTI-NETWORK CONNECTIVITY AND SERVICE DELIVERY ACROSS OPERATORS

To support use cases for new verticals and services, an operator might deploy a private network that serves only a set of the vertical industries and services. Meanwhile, the end-users within the private network might want to get access to other new services within the 5G network. To provide better user experience, the 5G system allows users to obtain service from multiple access network simultaneously on an on-demand basis. In addition, the 5G system can maintain service continuity with minimum service interruption when the serving network is changed from one network to another. An operator can have a variety of sharing business models and partnerships with other service providers to ensure its subscribers have access to multiple networks simultaneously. For a user with a single operator subscription, the use of multiple serving

network operators is under the control of the home operator. The user is served by the home operator as the first priority.

3.1.1.14 SUPPORT FOR VARIOUS ENHANCED V2X SCENARIOS

The 5G system includes requirements to support various enhanced V2X scenarios including:

- Vehicles platooning: The vehicles dynamically form a group travelling together. All the vehicles in the platoon receive periodic data from the leading vehicle so that minimum distance between vehicles can be maintained. All the vehicles that follow the leading vehicle can to be autonomously driven
- Advanced driving: Semi-automated or fully-automated driving are enabled. In this scenario, longer inter-vehicle distance is allowed. Each vehicle or RSU (Road Side Unit) shares data (obtained from its local sensors) with other vehicles in its proximity to coordinate driving routes and actions. Vehicles also share with each other their driving intention to avoid collision.
- Extended sensors: Local sensor data are exchanged among vehicles, RSUs, devices of pedestrians and V2X application servers. The exchanged data help to extend the perception of their environment and provide a more complete picture of the local situation to ensure safe driving.
- Remote driving: A remote driver or V2X application can operate a remote vehicle in the event that they cannot drive themselves or a remote un-manned vehicle is needed in dangerous environments. Cloud-computing-based driving can be used when the route variations are very limited, such as public transportation
- Vehicle QoS support: A V2X application can be notified of the estimated QoS change on time before actual change occurs. This will enable the 5G system to modify the V2X QoS parameters based on the actual service needs so that a constant user experience can be guaranteed

To support the previously listed V2X scenarios, the 5G system can control the communication range based on the characteristics of the transmitted message by a V2X UE, optimize the communication among V2X UEs belonging to the same group in proximity, support message transfer among a group of V2X UEs, support security and integrity of message transfers, support coordination of radio resources used for message transfers to maximize the spectral efficiency and ensure required reliability. For detailed requirements, please refer to Section 5.1 of 3GPP TS 22.186.

Performance requirements for enhanced V2X scenarios are defined in 3GPP Release 16 based on different levels of automation which are listed in Table 3.1.

Table 3.1. Level of automation.[113]

Level of Automation	0	1	2	3	4	5
Description	No Automation	Driver Assistance	Partial Automation	Conditional Automation	High Automation	Full Automation

[113] 3GPP TS 22.186 v16.1.0, *Enhancement of 3GPP support for V2X scenarios; Stage 1*. December 2018. The requirements listed here from 3GPP TS 22.186 are design targets in Rel-16. Some of these design targets may not be met by the Rel-16 specifications that are scheduled for completion by end of 2019.

Based on 3GPP Rel-16 specifications, the 5G system can support up to 5 UEs for a user group supporting V2X application. For Vehicle Platooning, the 5G system can support reliable Vehicle-to-Vehicle (V2V) communications between a specific V2X UE and up to 19 other V2X UEs and relative longitudinal position accuracy of less than 0.5 meters (m) for the Vehicle Platooning UEs in proximity. The detailed performance requirements are listed in Table 3.2.

Table 3.2. Performance Requirements for Vehicle Platooning.[114]

Communication scenario description		**Payload (Bytes)**	**Tx rate (Message / Sec)**	**Max end-to-end latency (ms)**	**Reliability (percent) (note 5)**	**Data rate (Mbps)**	**Min required communication range (meters) (note 6)**
Scenario	**Degree**						
Cooperative driving for vehicle platooning Information exchange between a group of UEs supporting V2X application	Lowest degree of automation	300-400 (note 2)	30	25	90		
	Low degree of automation	6500 (note 3)	50	20			350
	Highest degree of automation	50-1200 (note 4)	30	10	99.99		80
	High degree of automation			20		65 (note 3)	180
Reporting needed for platooning between UEs supporting V2X application and between a UE supporting V2X application and RSU	N/A	50-1200	2	500			
Information sharing for platooning between UE supporting V2X application and RSU	Lower degree of automation	6000 (note 3)	50	20			350
	Higher degree of automation			20		50 (note 3)	180
note 2: This value is applicable for both triggered and periodic transmission of data packets note 3: The data that is considered in this V2X scenario includes both cooperative manoeuvres and cooperative perception data that could be exchanged using two separate messages within the same period of time (for example, required latency 20 ms) note 4: This value does not include the security related messages component note 5: Sufficient reliability should be provided even for cells having no value in this table note 6: This is obtained considering UE speed of 130 km/h. All vehicles in a platoon are driving in the same direction							

In the case of advanced driving, the maximum allowable latency has to be small enough to ensure timely communication among vehicles in proximity. To coordinate among vehicles and to avoid collision, the maximum allowable end-to-end latency is 10 ms. In order to support emergent trajectory alignment among vehicles, only up to 3 ms end-to-end latency is allowed. In addition, vehicle platooning with the 5G system

[114] 3GPP TS 22.186 v16.1.0, *Enhancement of 3GPP support for V2X scenarios*; Stage 1. December 2018. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

has to ensure much higher reliability of the communication. The detailed performance requirements are listed in Table 3.3.

Table 3.3. Performance Requirements for Advanced Driving.[115]

Communication Scenario Description		**Payload (Bytes)**	**Tx rate (Message /Sec)**	**Max end-to-end latency (ms)**	**Reliability (percent) (NOTE 3)**	**Data rate (Mbps)**	**Min required Communication range (meters) (NOTE 4)**
Scenario	**Degree**						
Cooperative collision avoidance between UEs supporting V2X applications		2000 (NOTE 5)	100 (NOTE 5)	10	99.99	10 (NOTE 1)	
Information sharing for automated driving between UEs supporting V2X application	Lower degree of automation	6500 (NOTE 1)	10	100			700
	Higher degree of automation			100		53 (NOTE 1)	360
Information sharing for automated driving between UE supporting V2X application and RSU	Lower degree of automation	6000 (NOTE 1)	10	100			700
	Higher degree of automation			100		50 (NOTE 1)	360
Emergency trajectory alignment between UEs supporting V2X application		2000 (NOTE 5)		3	99.999	30	500
Intersection safety information between an RSU and UEs supporting V2X application		UL: 450	UL: 50			UL: 0. 25 DL: 50 (NOTE 2)	
Cooperative lane change between UEs supporting V2X applications	Lower degree of automation	300-400		25	90		
	Higher degree of automation	12000		10	99.99		
Video sharing between a UE supporting V2X application and a V2X application server						UL: 10	

NOTE 1: This includes both cooperative maneuvers and cooperative perception data that could be exchanged using two separate messages within the same period of time (for example, required latency 100 ms)
NOTE 2: This value is referring to a maximum number of 200 UEs. The value of 50 Mbps DL is applicable to broadcast or is the maximum aggregated bitrate of the all UEs for unicast
NOTE 3: Sufficient reliability should be provided even for cells having no values in this table
NOTE 4: This is obtained considering UE speed of 130km/h. Vehicles may move in different directions
NOTE 5: These values are based on calculations for cooperative maneuvers only

To support extended sensors, the minimum required communication range is much larger in some cases. The 5G system can provide up to 1km communication range when sensor information is shared among V2X UEs. The detailed performance requirements are listed in Table 3.4.

[115] 3GPP TS 22.186 v16.1.0, *Enhancement of 3GPP support for V2X scenarios; Stage 1*. December 2018.

Table 3.4. Performance Requirements for Extended Sensors.[116]

Communication scenario description		Payload (Bytes)	Tx rate (Message /Sec)	Max end-to-end latency (ms)	Reliability (percent)	Data rate (Mbps)	Minimum required communi-cation range (m)
Scenario	**Degree**						
Sensor information sharing between UEs supporting V2X application	Lower degree of automation	1600	10	100	99		1000
	Higher degree of automation			10	95	25 (NOTE 1)	
				3	99.999	50	200
				10	99.99	25	500
				50	99	10	1000
				10	99.99	1000	50
Video sharing between UEs supporting V2X application	Lower degree of automation			50	90	10	100
	Higher degree of automation			10	99.99	700	200
				10	99.99	90	400
NOTE 1: This is peak data rate NOTE 2: This is for imminent collision scenario							

For remote driving, the 5G system supports reliable and fast message exchange between a V2X UE and V2X application server. The moving speed of the V2X UEs is allowed to reach up to 250 km/h. The maximum allowable end-to-end latency is required to be 5 ms. In this case, the reliability has to be 99.999 percent. Table 3.5 summarizes the detailed requirements.

[116] 3GPP TS 22.186 v16.1.0, *Enhancement of 3GPP support for V2X scenarios*; Stage 1. December 2018. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

Table 3.5. Performance Requirements for Remote Driving.[117]

Communication Scenario Description	Max end-to-end latency (ms)	Reliability (percent)	Data rate (Mbps)
Information exchange between a UE supporting V2X application and a V2X Application Server	5	99.999	UL: 25 DL: 1

3.1.1.15 SUPPORT FOR UNIFIED ACCESS CONTROL

For massive IoT use cases, the 5G network is expected to support a huge number of IoT devices simultaneously. It is essential for the 5G system to provide a unified access control mechanism so that access to the network can be prevented when network congestion occurs. Based on operator's policy, the 5G system can prevent UEs from accessing the network using relevant barring parameters that depend on the Access Identity and Access Category. The Access Identities are configured at the UE based on Table 3.6. The Access Categories are defined based on the conditions of the UEs and the type of access attempt (Table 3.7). In unified access control, each access attempt is categorized into one or more of the Access Identities and one of the Access Categories. Based on the access control information applicable for the corresponding Access Identity and Access Category of the access attempt, the UE performs a test whether or not the actual access attempt can be accomplished. For detailed requirements for access control, refer to section 6.22 of TS 22.261.

[117] 3GPP TS 22.186 v16.1.0, *Enhancement of 3GPP support for V2X scenarios; Stage 1*. December 2018. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

Table 3.6. Access Identities of UEs.[118]

Access Identity Number	UE Configuration
0	UE is not configured with any parameters from this table
1 (NOTE 1)	UE is configured for Multimedia Priority Service (MPS)
2 (NOTE 2)	UE is configured for Mission Critical Service (MCS)
3-10	Reserved for future use
11 (NOTE 3)	Access Class 11 is configured in the UE
12 (NOTE 3)	Access Class 12 is configured in the UE
13 (NOTE 3)	Access Class 13 is configured in the UE
14 (NOTE 3)	Access Class 14 is configured in the UE
15 (NOTE 3)	Access Class 15 is configured in the UE

NOTE 1: Access Identity 1 is used by UEs configured for Multimedia Priority Service (MPS), in the PLMNs where the configuration is valid. The PLMNs where the configuration is valid are Home PLMN (HPLMN), PLMNs equivalent to HPLMN, and visited PLMNs of the home country
Access Identity 1 is also valid when the UE is explicitly authorized by the network based on specific configured PLMNs inside and outside the home country

NOTE 2: Access Identity 2 is used by UEs configured for MCS, in the PLMNs where the configuration is valid. The PLMNs where the configuration is valid are HPLMN or PLMNs equivalent to HPLMN and visited PLMNs of the home country. Access Identity 2 is also valid when the UE is explicitly authorized by the network based on specific configured PLMNs inside and outside the home country

NOTE 3: Access Identities 11 and 15 are valid in Home PLMN only if the Equivalent HPLMN (EHPLMN) list is not present or in any EHPLMN. Access Identities 12, 13 and 14 are valid in Home PLMN and visited PLMNs of home country only. For this purpose, the home country is defined as the country of the Mobile Country Code (MCC) part of the International Mobile Subscriber Identity (IMSI)

[118] 3GPP TS22.261 v16.7.0, *Service Requirements for the 5G System; Stage 1*. March 2019. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

Table 3.7. Access Categories.[119]

Access Category Number	Conditions Related to UE	Type of Access Attempt
0	All	Mobile Originated (MO) signaling resulting from paging
1 (NOTE 1)	UE is configured for delay tolerant service and subject to access control for Access Category 1, which is judged based on relation of UE's HPLMN and the selected PLMN	All except for Emergency
2	All	Emergency
3	All except for the conditions in Access Category 1.	MO signaling on Non-Access Stratum (NAS) level resulting from other than paging
4	All except for the conditions in Access Category 1	MMTEL voice (NOTE 3)
5	All except for the conditions in Access Category 1	Multimedia Telephony Service (MMTEL) video
6	All except for the conditions in Access Category 1	Short Message Service (SMS)
7	All except for the conditions in Access Category 1	MO data that do not belong to any other Access Categories (NOTE 4)
8	All except for the conditions in Access Category 1	MO signaling on Radio Resource Control (RRC) level resulting from other than paging
9-31		Reserved standardized Access Categories
32-63 (NOTE 2)	All	Based on operator classification

NOTE 1: The barring parameter for Access Category 1 is accompanied with information that define whether Access Category applies to UEs within one of the following categories:
a) UEs that are configured for delay tolerant service
b) UEs that are configured for delay tolerant service and are neither in their HPLMN nor in a PLMN that is equivalent to it
c) UEs that are configured for delay tolerant service and are neither in the PLMN listed as most preferred PLMN of the country where the UE is roaming in the operator-defined PLMN selector list on the SIM/USIM, nor in their HPLMN nor in a PLMN that is equivalent to their HPLMN
When a UE is configured for EAB, the UE is also configured for delay tolerant service. In case a UE is configured both for Extended Access Bearer (EAB) and for EAB override, when upper layer indicates to override Access Category 1, then Access Category 1 is not applicable

NOTE 2: When there is an Access Category based on operator classification and a standardized Access Category-- for both of which an access attempt can be categorized, and the standardized Access Category is neither 0 nor 2, the UE applies the Access Category based on operator classification. When there is an Access Category based on operator classification and a standardized Access Category-- for both of which an access attempt can be categorized, and the standardized Access Category is 0 or 2, the UE applies the standardized Access Category

NOTE 3: Includes Real-Time Text (RTT)

NOTE 4: Includes IP Multimedia Subsystem (IMS) Messaging

3.1.1.16 SUPPORT FOR QOS MONITORING

To provide satisfactory IoT services, such as URLLC, vertical industrial automation or V2X, the 5G network needs to guarantee their specific QoS requirements. However, under some circumstances (for example under strong interferences), the latency or packet loss might increase so that the network might not be able to meet those QoS requirements. In such cases, it is critical that the application and/or application server is notified in a timely manner. The 5G system can support real time end-to-end QoS monitoring/assurance

[119]3GPP TS22.261 v16.7.0, *Service Requirements for the 5G System; Stage 1*. March 2019. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

for mission critical services, such as URLLC, V2X and vertical automation. The network provides an interface to the application for QoS monitoring to initiate QoS monitoring, request QoS parameters, or request logging information. Different levels of granularity of QoS monitoring are also supported.

3.1.1.17 SUPPORT FOR CYBER-PHYSICAL CONTROL APPLICATIONS IN VERTICAL DOMAINS

One of very important perspectives of 5G IoT is support of cyber-physical control applications in vertical domains, such as industrial automation, smart grid, intelligent transportation systems, etcetera A Cyber-Physical System (CPS) includes both physical and computational components that interact with each other in various ways. It is monitored and controlled by a network server, and tightly integrated with internet. CPS applications in automation have different activity patterns, such as open-loop control, closed-loop control, sequence control and batch control. Communication for CPS applications also have different patterns, including periodic deterministic, a-periodic deterministic and non-deterministic communications. Deterministic communication has stable delay between transmission and reception of a message. In other words, the latency is bounded by a given threshold. For non-deterministic communication, the latency between transmission and reception of a message varies from case to case.

Even though many CPS use cases have similar KPI values, the 5G system will need to be sufficiently flexible to allow deployment configurations that can meet the needs of different verticals and different uses. To support CPS applications, the 5G system supports the following main features:

- Clock synchronization to manage time sensitive communications in industrial environments
- High accuracy positioning to accurately track IoT devices as well as mobile assets in factories
- Very high communication service availability and reliability
- Very low end-to-end latency in many use cases

3GPP SA1 working group has defined key target performance for various automation use cases in Table 3.8.

Table 3.8. 3GPP Performance Requirements for Industrial Automation.[120]

Service	End-to-End Latency	Jitter	Survival time	Availability	Experienced Data Rate	Connection Density
Factory automation - Motion Control	1 ms	1 µs	0 ms	99,9999 percent	10 Mbps	100 000/km^2
Factory automation	10 ms	100 µs	0 ms	99,99 percent	10 Mbps	100 000/km^2
Process automation – Remote Control	50 ms	20 ms	100 ms	99,9999 percent	100 Mbps	1 000/km^2
Process automation Monitoring	50 ms	20 ms	100 ms	99,9 percent	1 Mbps	10 000/km^2

More automation performance values are provided in Table 3.9 by ZVEI electrical group.

[120] 3GPP Technical Specification 22.261, *Service requirements for the 5G system Stage 1*. February 2017. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

Table 3.9. ZVEI Performance Requirements for Industrial Automation.[121]

Use case (high level)		Availability	Cycle time	Typical payload size	# of devices	Typical service area
Motion control	Printing machine	>99.9999%	< 2 ms	20 bytes	>100	100 m x 100 m x 30 m
	Machine tool	>99.9999%	< 0.5 ms	50 bytes	~20	15 m x 15 m x 3 m
	Packaging machine	>99.9999%	< 1 ms	40 bytes	~50	10 m x 5 m x 3 m
Mobile robots	Cooperative motion control	>99.9999%	1 ms	40-250 bytes	100	< 1 km²
	Video-operated remote control	>99.9999%	10 – 100 ms	15 – 150 kbytes	100	< 1 km²
Mobile control panels with safety functions	Assembly robots or milling machines	>99.9999%	4-8 ms	40-250 bytes	4	10 m x 10 m
	Mobile cranes	>99.9999%	12 ms	40-250 bytes	2	40 m x 60 m
Process automation (process monitoring)		>99.99%	> 50 ms	Varies	10000 devices per km²	

Source: ZVEI

To support cyber-physical system applications using Ethernet, the 5G system provides an Ethernet transport service. In this case, the network has an infrastructure dedicated to high-performance Ethernet applications. In addition to the general requirements to support 5G LAN-type services listed in Section 6.24 in TS 22.261, 3GPP defined detailed requirements specifically for Cyber-Physical System applications using Ethernet. The Ethernet transport service supports routing based on information extracted from the Ethernet header information created in 802.1Qbv. The 5G system supports clock synchronization defined by IEEE 802.1AS across 5G-based Ethernet links with Protocol Data Unit (PDU)-session type Ethernet, as well as other Ethernet transports such as wired and optical. The accuracy of the clock synchronization is better than 1µs. Enhancements for time-sensitive networking as defined by IEEE 802.1Q, such as time-aware scheduling with absolute cyclic time boundaries are supported. Cyclic time boundaries are configurable for both downlink and uplink. The 5G system also supports co-existence of real time traffic following a time-aware schedule and lower priority traffic as long as the lower priority traffic does not degrade the performance of the real time traffic.

3.1.1.18 SUPPORT FOR ETHERNET TRANSPORT SERVICES

As mentioned previously, Ethernet transport services support Cyber-Physical System applications on Ethernet. Routing of non-IP packet (for example, Ethernet frame) is supported efficiently for private communication between UEs within a 5G LAN-type service. The network provides the required QoS for non-IP packets. The 5G system also supports routing based on information extracted from Virtual LAN (VLAN) ID, from Bridge Protocol Data Units crated in the Ethernet network based on Spanning Tree Protocol. Traffic filtering and prioritization are also supported.[122]

3.1.1.19 SUPPORT FOR NON-PUBLIC NETWORKS

[121] *5G for Connected Industries and Automation*, 5G ACIA. February 2019. www.5g-acia.org.
[122] 3GPP TS22.261 V 16.7.0, *Service Requirements for the 5G System; Stage 1*. March 2019. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

Many IoT applications will be supported in non-public networks. They may be deployed as completely stand-alone networks, or may be hosted by a PLMN, or may be offered as a slice of a PLMN. The 5G system supports both physical and virtual non-public networks. Depending on operator policies and regulatory requirements, the 5G system may support non-public network subscribers' access to subscribed PLMN services via the non-public network. Non-public network subscribers should also be able to get access to selected non-public network services via PLMN. Seamless service continuity between the non-public network and PLMN is supported. In order to access a PLMN service, a non-public network subscriber will have a service subscription using 3GPP identifiers and credentials provided or accepted by a PLMN. A UE will be able to identify and select a non-public network. The 5G system will support a large number of Identifiers for non-public networks to avoid possible collision between assigned Identifiers. A UE with a subscription to a non-public network will be prevented from automatically selecting and attaching to a PLMN or another non-public network that it is not authorized to select. On the other hand, a UE with a subscription to a PLMN will also be prevented from automatically selecting and attaching to a non-public network that it is not authorized to select. If an IoT application is originally hosted by a PLMN, and then later the host is changed to another PLMN, the network selection information stored in the IoT devices of the non-public network will not be changed.

3.1.1.20 SUPPORT FOR 5G LAN-TYPE SERVICE

There are multiple IoT applications in residential areas, offices, enterprises and factories, where 5G will need to provide services with similar functionalities to Local Area Networks (LANs) and Virtual Public Networks (VPNs) and with improved performance and security. 5G LAN-type service is supported in both unlicensed and licensed spectrum.

An IoT device is able to select a 5G-LAN Virtual Network (5G-LAN-VN) to which it belongs for private communications. The 5G system is able to support a 5G-LAN-VN with a large number of members (up to tens of thousands). On-demand connection of UE to UE, multicast and broadcast private communications are supported between members of the same 5G-LAN-VN. Consistent Quality of Experience (QoE) is provided to all members. For flexibility, the operator can create, manage and remove a 5G LAN-VN based on its needs. For privacy, the 5G system can prevent the sharing of a UE's identification information on private communication among LAN members. Traffic scenarios found in an industrial setting are supported for 5G LAN-type services.

5G LAN-type service can be provided for authorized users using both direct and indirect network connection (connection through a relay UE). Depending on operator's policy, the 5G network provides suitable APIs to allow a trusted third party to create or remove a 5G LAN-VN.

3.1.1.21 SUPPORT FOR POSITIONING SERVICES

Many vertical IoT applications such as V2X, UAVs or industrial automatic control, require very high positioning accuracy. It is important for the 5G system to provide 5G positioning services in compliance with regulatory requirements. The 5G system provides different 5G positioning services to provide absolute and relative positioning. Both single and hybrid positioning methods are supported. Hybrid positioning methods include both the 3GPP positioning technologies and non-3GPP positioning technologies, such as, Global Navigation Satellite System (GNSS), network-based assisted GNSS, Terrestrial Beacon Systems, dead-reckoning sensors, or WLAN/Bluetooth-based positioning.

5G positioning services is provided in case of roaming. Even when a UE is outside the coverage of 3GPP RAT-dependent technologies but within 5G positioning service area (for example, within the coverage of

satellite access), the 5G system can determine its position. The positioning-related data can be logged and made available to an application, or an application server. The data is then managed in compliance with traceability requirements as well as requirements of authentication and security. A UE can provide the 5G system with its position-related data periodically or on request. The transmission interval of the UEs' position-related data is configurable based on different location models and performance criteria, such as power consumption or service latency. Depending on different IoT applications, the period of data updating can range from 0.1 second (s) to one set of positioning-related data every month. For more detailed requirements, please refer to Section 6.27 of TS 22.261.

3.1.2 RELEASE 17

With Rel-16 Phase 1 complete in 4Q18, 3GPP SA1 has begun work on Rel-17 requirements. Most of the new study items are driven by vertical/enterprise needs for IOT support. Many of them build directly on the baseline of IOT support in Rel-16. The active studies for additional IOT enhancements include the following:

- TR 22.826, Feasibility Study on Communication Services for Critical Medical Applications (FS_CMED) – using 5G for communications for critical medical scenarios such as assisted surgery and monitoring health status
- TR 22.827, Feasibility Study on Audio-Visual Service Production (FS_AVPROD) – enhancements to URLLC and TSN requirements specifically for A/V use cases
- TR 22.829, Study on enhancement for UAVs (FS_EAV) – additional enhancements for UAV control and monitoring
- TR 22.836, Study on Asset Tracking Use Cases (FS_5G_ATRAC) – using 5G for communications with moving assets such as shipping containers, wagons, and pallets
- TR 22.832, Study on enhancements for Cyber-Physical Control applications in vertical domains (FS_eCAV) – additional enhancements for communications for industrial automation

The studies mentioned consider specific use cases in each enterprise domain that may generate new requirements for 5G systems. In some cases, this results in potential new sets of key performance indicators (KPIs) such as the specific combination of availability, reliability, latency, number of UEs served, and service area needed to support the use case. In other cases, this results in potential requirements for specific functionality in the network or new Application Programming Interfaces (APIs) to support the use case. The studies each include a gap analysis with previous 5G work to ensure that the potential requirements are not already covered by requirements in previous releases.

Other Rel-17 studies are underway which may lead to requirements useful for IOT while not specifically focused on IOT applications:

- TR 22.842, Study on Network Controlled Interactive Service in 5GS (FS_NCIS)
- TR 22.866, Study on enhanced Relays for Energy eFficiency and Extensive Coverage (FS_REFEC) – extending the Proximity Services (ProSe) UE-to-Network relay to include multi-hop relay UEs.

3.2 VERTICAL REQUIREMENTS

Clearly, apart from the generic requirements common to all IoT services and presented in the previous sections, the specific needs for each of the wide variety of services under consideration may differ.

From an operator point of view, there are different requirements and estimations, depending upon the operator role and ambition. As an example, in the METIS II project the "Massive distribution of sensors and actuators" use case, the following views on requirements for future 5G developments, as shown in Table 3.10 are provided.

Table 3.10. Refined Scenarios and Requirements, Consolidated Use Cases and Qualitative Techno-economic Feasibility Assessment.[123]

Availability	99.9 percent
Device density Traffic volume per device	1 000 000 devices/km[124] 125 bytes message per second
Battery life	10 years (assuming 5 Watts-hour battery and restricted traffic model)

Due to the high number of possible IoT services, the variety of uses cases with different requirements is tremendous. In Table 3.11, some of the most relevant use cases, in terms of a clear business case and their likelihood to be exploited in the next few years, are summarized.

Table 3.11. Prospective Use Cases for IoT.

Sector	Use Case	Top Requirements
Industry	High Volume (for example, mining)	Range, Coverage, Reliability, Cost
Agriculture	Dynamic (for example, animal tracking)	Battery, Range, Coverage, Reliability, Cost
	Static (for example, irrigation of fields)	Battery, Range, Coverage, Reliability, Cost
Utilities	Powered (for example, Electricity)	Indoor, Service Level Agreement (SLA), Reliability
	Not Powered (for example, Water/Gas)	Indoor, SLA, Reliability
Logístics	Management & Tracking (for example, Fleet)	Easy Install., Mobility, Coverage, Cost
	Basic Monitoring (for example, shipment conditions, warehouse)	Battery, Easy Install., Mobility, Coverage, Cost
Smart Cities	Dynamic Systems (for example, Traffic Management)	SLA, Coverage, Reliability
	Basic Sensoring (for example, air pollution)	SLA, Coverage, Reliability

[123] METIS II Deliverable D1.1 2016-01-31.
[124] *LTE Progress Leading to the 5G Massive Internet of Things*, 5G America White Paper, November 2017.

Payments	Total Payment Volume (TPV)	Indoor, Interoperability, SLA, Reliability
	Fraud Detection	Indoor, Interoperability, SLA, Reliability
Wearables (incl. e-Health)	Continuous Tracking (for example, Diabetes)	Indoor, Battery, Mobility, SLA, Coverage, Reliability
	Spot Tracking (for example, steps tracking)	Battery, Easy Install., Mobility
Security	High Volume (for example, video)	Indoor, Throughput, Security, SLA, Reliability
	Low Volume (for example, presence detection)	Indoor, Security, SLA, Reliability
Connected Cars	Integrated solution (for example, traffic management)	Easy Install., Mobility, Coverage, Cost
	Basic Monitoring (for example, location)	Easy Install., Mobility, Coverage, Cost
Buildings (incl. Home)	Complex Solution (for example, energy management)	Indoor, Security, SLA, Reliability
	Basic Solution (for example, presence/ air pollution)	Indoor, Security, SLA, Reliability
IoT Complex Systems	Autonomous Car or Drones Ecosystems	Battery, Security, Range, SLA, Coverage, Reliability

Depending upon the specific service and the values, different sets of requirements should be considered to meet specific IoT service needs, such as:

- Traffic patterns (throughput and active cycles)
- Identity/Security needs
- Ease of Installation
- Mobility
- Service Level Agreement (SLA)
- Reliability
- Possible sector regulations
- Analytics and charging needs

Requirements for the large variety of IoT applications are vastly different. 3GPP is working to develop a set of global standards to ensure that the challenges are successfully met with values that will accommodate the myriad of services and applications that our connected future will present.

An enterprise with a strong desire for the security of its data may prefer a private network for its IoT needs. Such a network can leverage 5G technology but will operate separately from operators' public cellular networks, providing access only to authorized devices. In some cases, operators may still deploy and operate such private networks as a service to the enterprises as the latter often lack the experience and expertise to do so. Operators with spectrum to spare can opt to deploy a private network in one of their unutilized carriers. Alternately, they can also deploy the private network in shared or unlicensed spectrum.

In addition to the requirements listed in Tables 3.10 and 3.11, a private network will also need to support the following:

- Independence from the public cellular network
- Dedicated network equipment in the enterprise's premise with ubiquitous coverage
- QoS control for different equipment
- High degree of reliability

3.2.1 APPLICATION AREAS AND USE CASES OF CYBER-PHYSICAL APPLICATIONS IN VERTICAL DOMAINS

The 5G system is expected to support Cyber-Physical Control applications in vertical domains, which are mostly related to industrial automation and robotics. 3GPP defined five distinct areas of applications as follows:[125]

- Factory automation: Comprises automated control, monitoring and optimization of processes and workflows within a factory. Example use cases include motional control, control-to-control, mobile robots and massive wireless sensor networks. Communication services need to fulfill stringent requirements in terms of latency, communication service availability and determinism.
- Process automation: Deals with production control and handling of substances such as chemicals, food and beverages. Example use cases include mobile robot, massive wireless sensor networks, closed-loop process control, process monitoring and plant asset management. Communication services need to meet stringent requirements in terms of latency and determinism.
- Human-machine interfaces (HMIs) and production IT: HMIs include many diverse devices for interaction between people and production systems, while production IT includes Manufacturing Execution System (MES) and Enterprise Resource Planning (ERP) systems for process monitoring. Example use cases include mobile control panels and Augmented Reality (AR) and Virtual Reality (VR) systems. Communication services need to meet stringent requirements in terms of latency, communication availability and determinism.
- Logistics and warehousing: Organize and control the industrial production flow and storage of materials and goods. Example of use cases include control-to-control and mobile robots. Communication services need to meet stringent requirements in terms of latency, communication availability and determinism.
- Monitoring and predictive maintenance: Handles monitoring of certain processes and/or assets but does not have immediate impact on the processes themselves. Example of use cases include massive wireless sensor networks, remote access and maintenance.

As can be seen, for each of these application areas, a multitude of potential use cases exists. Some use cases include multiple application areas. For example, mobile robots include applications of factory automation, process automation as well as logistics and warehousing. Table 3.12 summarizes the mapping of application areas and use cases.

Table 3.12. Mapping of the Vertical IoT Use Cases to Application Areas.[126]

	Factory Automation	Process Automation	HMIs and Production IT	Logistics and Warehousing	Monitoring and Maintenance
Motion Control	X	X			
Control-to-Control	X			X	

[125] 3GPP TS 22.104 V16.1.0, Service requirements for cyber-physical control applications in vertical domains; Stage 1. March 2019. The requirements listed here from 3GPP TS 22.104 are design targets in Rel-16. Some of these design targets may not be met by the Rel-16 specifications that are scheduled for completion by end of 2019.
[126] *Ibid.*

Mobile Control Panels with Safety			X		
Mobile Robots	X			X	
Remote Access and Maintenance					X
Augmented Reality			X		
Closed-Loop Process Control		X	X		
Process Monitoring		X			
Plant Asset Management		X	X	X	

Motion control: Responsible for controlling moving and/or rotating parts of machines in a well-defined manner.
Control-to-control: Communication between different industrial controllers
Mobile robots: Programmable machine able to execute multiple operations. Autonomous guided vehicle is an example
Closed-loop process control: Closed-loop in process automation
Plant asset management: Timely diagnosis and maintenance of the plant components

All the listed applications follow certain activity patterns, which are either open-loop or closed-loop control. Communication services supporting those applications also follow certain patterns, which are periodic deterministic, aperiodic deterministic and non-deterministic communication. In the case of periodic deterministic communication, transmission interval is repeated and the delay between transmission and receipt of a message is stable (bounded by a given threshold). In case of non-deterministic communication, the delay between transmission and receipt of a message varies from time to time.

In the case of open-loop control, one or many messages are sent to an actuator and no feedback is expected. The output of the influenced process is assumed to be predetermined and within an acceptable range. Open-loop control works when the environment impacts on the process and the actuators are negligible. The messages can be sent in a periodic or an aperiodic pattern, however the communication pattern has to be deterministic so that the control process can work without an activity response from the receiver. Use cases involving monitoring and predictive maintenance usually use open-control.

Closed-loop can enable the control processes even if the environment impacts on the process are non-negligible or the performance of the actuator changes over time. During the closed-loop control process, an actuator senses the process output and then feeds measurements back to the controller. Closed-loop process produces both periodic and aperiodic communication patterns. On the other hand, this kind of process is often used for continuous control processes with tight time-control limits. In this case, the process typically relies on periodic communication patterns. The communication has to be deterministic. Use cases for factory automation and process automation usually use closed-loop control.

Logging of device states, measurements, etcetera, for maintenance purposes typically have aperiodic communication patterns. Communication patterns can be non-deterministic if the transmitted logging information can be time stamped.

Depending on activity and communication patterns, 3GPP defined the performance requirements for different vertical applications, which will be summarized in the following sections. In most of the vertical applications, clock synchronization is needed for managing time sensitive communications in an industrial environment. High accuracy positioning is also a very important requirement for lots of vertical use cases for industrial control. In addition to clock synchronization and high positioning accuracy, very high communication service availability is required for Cyber-Physical Application, especially for applications with deterministic traffic. To meet the requirement of the very high communication availability, the 5G

system meets the stringent requirements of system latency, service survival time and system reliability. If the message transfer time exceeds the maximum allowable latency and the message is not received on time, the communication service is considered to be unavailable. An example of the relationship between communication service availability and system reliability of logical links is illustrated in Table 3.13.

Table 3.13. Relationship between Communication Service Availability and Reliability.[127]

Communication service availability	Reliability (as defined in TS 22.261)
99,999 %	99,9 %
99,999999 %	99,99 %
99,9999999 %	99,999 %
99,9999999999 %	99,9999 %
99,999999999999 %	99,99999 %

3.2.1.1 PERFORMANCE REQUIREMENTS FOR PERIODIC DETERMINISTIC COMMUNICATION

In case of periodic deterministic communication, a transmission occurs every transfer interval. This type of communication has very stringent requirements on both end-to-end latency and communication service availability. For example, for control-to-control procedure, the required communication service availability has to be 99.9999 percent - 99.999999 percent. The maximum allowable latency is only 500 µs in the case of motion control. The latency requirement can be a bit more relaxed in the case of process monitoring and plant asset management. The maximum allowable end-to-end latency can be 100 ms – 60 s. These requirements have to be met under quite high UE speed, the maximum UE speed can be 75 km/h. In most use cases except video operated remote control, the message size is small, between 20 and 250 bytes. The use case of video operated remote control has the biggest message size, which is 15 – 250 kbytes, which allows relatively larger latency (10 – 100 ms). Table 3.14 presents the detailed performance requirements for different use cases.

Table 3.14. Periodic Deterministic Communication Service Performance Requirements.[128]

Use Case	Performance Matrix		Target Value			Note
Motion Control	Communication service availability	Target value	99.999 percent - 99.99999 percent			One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement. All communication includes 1 wireless link (UE to network node or network node to UE)
		Mean time between failures	~ 10 years			
	Maximum End-to-end latency		< transfer interval			
	User experienced data rate		-			
	Message size [bytes]		50	40	20	
	Transfer interval [ms]		0.5	1	2	
	Survival time		0.5			
	UE speed		≤ 75 km/h			
	# of UEs		≤ 20	≤ 50	≤ 100	
	Service area [length x width x height]		50 m x 10 m x 10 m			

[127] 3GPP TS 22.104 V16.1.0, *Service requirements for cyber-physical control applications in vertical domains; Stage 1*. March 2019. The requirements listed here from 3GPP TS 22.104 are design targets in Rel-16. Some of these design targets may not be met by the Rel-16 specifications that are scheduled for completion by end of 2019.
[128] *Ibid.*

Electrical Distribution **(Distributed Automated Switching for Isolation and Service Restoration)**	Communication service availability	Target value	99.9999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement Communication includes two wireless links (UE to UE)
		Mean time between failures	-	
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		1 kbps (steady state) 1.5 Mbps (fault case)	
	Message size		< 1500 bytes	
	Transfer interval		< 60 s (steady state) ≥ 1 ms (fault case)	
	Survival time		TBD	
	UE speed		stationary	
	# of UEs		20	
	Service area [length x width]		30 km x 20 km	
Control-to-Control in Motion Control	Communication service availability	Target value	99.9999 percent - 99.999999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement Communication may include two wireless links (UE to UE)
		Mean time between failures	~ 10 years	
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		-	
	Message size		1 kbytes	
	Transfer interval		≤ 10 ms	
	Survival time		10 ms	
	UE speed		-	
	# of UEs		5 - 10	
	Service area [length x width x height]		100 m x 30 m x 10 m	
Mobile Robots	Communication service availability	Target value	>99.9999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement All communication includes 1 wireless link (UE to network node or network node to UE) This category covers different transfer intervals for different similar use cases with target values of 1ms, 1 – 10 ms, and 10 – 50 ms. The transfer interval deviates around its target value by < ± 25 percent
		Mean time between failures	~ 10 years	
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		-	
	Message size		40 – 250 bytes	
	Transfer interval		1 ms – 50 ms	
	Survival time		Transfer interval	
	UE speed		≤ 50 km/h	
	# of UEs		≤ 100	
	Service area [length x width]		≤ 1 km^2	
Mobile Robots **(Video-operated remote control)**	Communication service availability	Target value	>99.9999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement All communication includes 1 wireless link (UE to network node or network node to UE) The transfer interval deviates around its target value by < ± 25 percent
		Mean time between failures	~ 1 year	
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		-	
	Message size		15 – 250 kbytes	
	Transfer interval		10 – 100 ms	
	Survival time		Transfer interval	
	UE speed		≤ 50 km/h	
	# of UEs		≤ 100	
	Service area [length x width]		≤ 1 km^2	
		Target value	>99.9999 percent	

Mobile Robots	Communication service availability	Mean time between failures	~1 year	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement All communication includes 1 wireless link (UE to network node or network node to UE) The transfer interval deviates around its target value by < ± 25 percent
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		-	
	Message size		10 – 250 bytes	
	Transfer interval		50 – 500 ms	
	Survival time		Transfer interval	
	UE speed		≤ 50 km/h	
	# of UEs		≤ 100	
	Service area [length x width]		≤ 1 km^2	
Mobile Control Panels **(Remote Control of for example, Assembly Robots, Milling Machines**	Communication service availability	Target value	99.9999 percent - 99.999999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement Communication may include two wireless links (UE to UE) The transfer interval deviates around its target value by ± 25 percent
		Mean time between failures	~ 1 month	
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		-	
	Message size		40 – 250 bytes	
	Transfer interval		4 ms – 8 ms	
	Survival time		Transfer interval	
	UE speed		< 8 km/h	
	# of UEs		TBD	
	Service area [length x width x height]		50 m x 10 m x 4 m	
Mobile Control Panels **(Remote Control of for example, mobile cranes, mobile pumps, fixed portal cranes**	Communication service availability	Target value	99.9999 percent - 99.999999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement Communication may include two wireless links (UE to UE) The transfer interval deviates around its target value by ± 25 percent
		Mean time between failures	~ 1 year	
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		-	
	Message size		40 – 250 bytes	
	Transfer interval		< 12 ms	
	Survival time		12 ms	
	UE speed		< 8 km/h	
	# of UEs		TBD	
	Service area [length x width]		Typically, 40 m x 60 m; Maximum, 200 m x 300 m	
Process Automation **(Closed Loop Control)**	Communication service availability	Target value	99.9999 percent - 99.999999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement All communication includes 1 wireless link (UE to network node or network node to UE) The transfer interval deviates around its target value by < ± 5 percent
		Mean time between failures	≥ 1 year	
	Maximum End-to-end latency		< transfer interval	
	User experienced data rate		-	
	Message size		20 bytes	
	Transfer interval		≥ 10 ms	
	Survival time		0	
	UE speed		Typically stationary	
	# of UEs		10 - 20	
	Service area [length x width x height]		≤ 100 m x 100 m x 50 m	

Primary Frequency Control	Communication service availability	Target value	99.999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement Communication may include two wireless links (UE to UE)
		Mean time between failures	TBD	
	Maximum End-to-end latency		~ 50 ms	
	User experienced data rate		-	
	Message size		~ 100 bytes	
	Transfer interval		~ 50 ms	
	Survival time		TBD	
	UE speed		stationary	
	# of UEs		≤ 100 000	
	Service area [length x width]		Several km^2 to 100000 km^2	
Distributed Voltage Control	Communication service availability	Target value	99.9999 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement Communication may include two wireless links (UE to UE)
		Mean time between failures	TBD	
	Maximum End-to-end latency		~100 ms	
	User experienced data rate		-	
	Message size		~100 bytes	
	Transfer interval		~200 ms	
	Survival time		TBD	
	UE speed		stationary	
	# of UEs		≤ 100 000	
	Service area [length x width]		several km^2 up to 100 000 km^2	
Process Monitoring, **Plant Asset Management**	Communication service availability	Target value	99.99 percent	One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement. All communication includes 1 wireless link (UE to network node or network node to UE)

3.2.1.2 PERFORMANCE REQUIREMENTS FOR APERIODIC DETERMINISTIC COMMUNICATION

Aperiodic deterministic communication does not have a pre-set sending time, but still has very stringent requirements on latency and communication service availability. An aperiodic transmission can be triggered instantaneously by an event, such as:

- Process event that comes from the process when thresholds are exceeded or fallen below, for example, temperature or pressure, and etcetera
- Diagnostic event that indicates malfunctions of an automation device or module, for example, power supply defective, short circuit, or two high temperature, and etcetera
- Maintenance event based on information that indicates necessary maintenance work to prevent the failure of an automation device

After an event is triggered and an alarm signal has been received by the application, an acknowledgement is usually sent within a short period of time. If no acknowledge is received after a preset monitoring time, either the alarm will be re-sent, or some failure response action is started.

Compared to periodic deterministic use cases, the maximum end-to-end latency requirements are more relaxed in the case of aperiodic deterministic use cases. The target value ranges from 10 to 50 ms. Among

all the listed aperiodic deterministic use cases, AR (Augmented Reality) has the tightest latency requirement, which is < 10 ms. Table 3.15 presents the detailed requirements.

Table 3.15. Aperiodic Deterministic Communication Service Performance Requirements.[129]

Use Case	Performance Matrix		Target Value	Note
Mobile Robots Video Streaming	Communication service availability	Target value	>99.9999 percent	All communication includes 1 wireless link (UE to network node or network node to UE)
		Mean time between failures	~ 1 week	
	Maximum end-to-end latency		10 ms	
	User experienced data rate		>10 Mb/s	
	Message size [bytes]			
	Survival time			
	UE speed		≤ 50 km/h	
	# of UEs		≤ 100	
	Service area [length x width]		≤ 1 km^2	
Mobile Control Panels **Parallel Data Transmission**	Communication service availability	Target value	99.9999 percent - 99.999999 percent	All communication includes 1 wireless link (UE to network node or network node to UE)
		Mean time between failures	~ 1 month	
	Maximum end-to-end latency		< 30 ms	
	User experienced data rate		>5Mb/s	
	Message size			
	Survival time			
	UE speed		< 8 km/h	
	# of UEs		TBD	
	Service area [length x width]		TBD	
Smart Grid ms-Level Precise Load Control	Communication service availability	Target value	99.9999 percent	All communication includes 1 wireless link (UE to network node or network node to UE)
		Mean time between failures	-	
	Maximum end-to-end latency		< 50 ms	
	User experienced data rate		0.59 – 28 kb/s	
	Message size		<100 bytes	
	Survival time			
	UE speed		stationary	
	# of UEs		10 – 100 per km^2	
	Service area [length x width x height]		TBD	
Augmented Reality	Communication service availability	Target value	>99.99 percent	All communication includes 1 wireless link (UE to network node or network node to UE)

[129] 3GPP TS 22.104 V16.1.0, *Service requirements for cyber-physical control applications in vertical domains; Stage 1.* March 2019. The requirements listed here from 3GPP TS 22.104 are design targets in Rel-16. Some of these design targets may not be met by the Rel-16 specifications that are scheduled for completion by end of 2019.

Bi-directional Transmission to Image Processing				

3.2.1.3 PERFORMANCE REQUIREMENTS FOR NON-DETERMINISTIC COMMUNICATION

In the case of non-deterministic communication, the delay between the transmission and receipt of a message varies from time to time, and cannot be bounded by a threshold. This includes periodic or aperiodic non-real-time traffic. This type of traffic is more tolerant to the latency. Compared to deterministic traffic, the required communication service availability is also lower. Table 3.16 shows the performance requirements of non-deterministic traffic.

Table 3.16. Non-deterministic Communication Service Performance Requirements.[130]

Use Case	Performance Matrix		Target Value
Motion Control Software Updates	Communication service availability	Target value	
		Mean time between failures	~ 1 month
	User experienced data rate		>1 Mb/s
	UE speed		≤ 75 km/h
	# of UEs		≤ 100
	Service area [length x width]		50 m x 10 m x 10 m
Mobile Robots Real-Time Video Stream	Communication service availability	Target value	
		Mean time between failures	
	User experienced data rate		>10 Mb/s
	UE speed		≤ 50 km/h
	# of UEs		≤ 100
	Service area [length x width]		≤ 1 km^2

3.2.1.4 PERFORMANCE REQUIREMENTS FOR MIXED TRAFFIC

Mixed traffic cannot be categorized into the above listed communication patterns, so the performance requirements are listed separately, which is listed in Table 3.17. As can be seen, wind power plant requires very high communication availability and very low packet rate for communication.

Table 3.17. Mixed Traffic Communication Service Performance Requirements.[131]

Use Case	Performance Matrix		Target Value	Note
Wind Power Plant	Communication service availability	Target value	>99.9999999 percent	All communication includes 1 wireless link (UE to network node or network node to UE)
		Mean time between failures	~ 10 years	
	Maximum end-to-end latency		16 ms	
	Packet error ratio		< 10^{-9}	
	UE speed		stationary	

[130] 3GPP TS 22.104 V16.1.0, *Service requirements for cyber-physical control applications in vertical domains*; *Stage 1,* March 2019. The requirements listed here from 3GPP TS 22.104 are design targets in Rel-16. Some of these design targets may not be met by the Rel-16 specifications that are scheduled for completion by end of 2019.

[131] *Ibid.*

	# of UE	< 1000	
	Service area [length x width]	several km^2	

3.2.1.5 CLOCK SYNCHRONIZATION REQUIREMENTS

Clock synchronization is very important for many use cases in vertical domain. The detailed requirements were given in 3GPP TS22.104 Section 5.69. To support clock synchronization, data processing and transmission is done in accordance with IEEE 1588v2 (Precision Time Protocol) to support third-party applications that use this protocol. The 5G system supports a mechanism to synchronize the user-specific time clock of UEs with a global clock and/ or working clock. To reflect the actual complexity of a real factory, up to 32 working block domains are supported in the network. The requirement on the synchronization precision for industrial control is very high. The global time domain provides time synchronization with precision of 1 µs. The working block domains also provide similar synchronization precision of ≤ 1 µs, where the precision is defined between the sync mater and any device of the clock domain. The synchronization precision level is slightly lower (< 10 µs) in case of high data rate video streaming. Table 3.18 summarizes the performance requirements of clock synchronization.

Table 3.18. Clock Synchronization Service Performance Requirements.[132]

User-specific clock synchronicity accuracy level	# of devices in one Communication group for clock synchronisation	Clock synchronicity requirement	Service area	Scenario
1	Up to 300 UEs	< 1 µs	≤ 100 m x 100 m	Motion control Control-to-control communication for industrial controller
2	Up to 10 UEs	< 10 µs	≤ 2500 m^2	High data rate video streaming
3	Up to 100 UEs	< 1 µs	< 20 km^2	Smart Grid: synchronicity between PMUs

3.2.1.6 POSITIONING PERFORMANCE REQUIREMENTS

In all the use cases of industrial automation and process control, tracking of IoT devices and mobile assets is becoming increasingly important for process improvement and flexible operation in industrial environments. Positioning with high accuracy is therefore essential for Factories of the Future. For different service area, environment and UE speeds, there are different performance requirements in terms of positioning accuracy. 3GPP defined 7 positioning service levels based on different scenarios[133]. Depending on vertical IoT use cases, the required positioning service levels range from level 2 to level 7. For example, mobile control panels with safety functions within non-danger zones is required to have positioning service level of 2, where the horizontal accuracy is within 5m, the availability of positioning service is 90 percent and the latency for UE positioning estimation is less than 5s. AR (Augmented Reality) in smart factories is required to have positioning service level of 4, where the horizontal accuracy is within

[132] 3GPP TS 22.104 V16.1.0, Service requirements for cyber-physical control applications in vertical domains; Stage 1. March 2019. The requirements listed here from 3GPP TS 22.104 are design targets in Rel-16. Some of these design targets may not be met by the Rel-16 specifications that are scheduled for completion by end of 2019.

[133] 3GPP TS 22.261 V16.7.0, Service requirements for the 5G system; Stage 1. March 2019. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

1m, the availability of positioning service is 99 percent and the latency for UE positioning estimation is less than 15 ms. Table 3.19 summarizes the positioning performance requirements for different vertical IoT scenarios.

Table 3.19 Performance Requirements for Positioning Accuracy.[134]

Scenario	Horizontal accuracy	Availability	Heading	Latency for UE position estimation	UE Speed	Corresponding Positioning Service Level in TS 22.261
Mobile control panels with safety functions (non-danger zones)	< 5 m	90 percent	N/A	< 5 s	N/A	Service Level 2
Process automation – plant asset management	< 1 m	90 percent	N/A	< 2 s	< 30 km/h	Service Level 3
Flexible, modular assembly area in smart factories (for tracking of tools at the work-place location)	< 1 m (relative positioning)	99 percent	N/A	1 s	< 30 km/h	Service Level 3
Augmented reality in smart factories	< 1 m	99 percent	< 0,17 rad	< 15 ms	< 10 km/h	Service Level 4
Mobile control panels with safety functions in smart factories (within factory danger zones)	< 1 m	99.9 percent	< 0,54 rad	< 1 s	N/A	Service Level 4
Flexible, modular assembly area in smart factories (for autonomous vehicles, only for monitoring proposes)	< 50 cm	99 percent	N/A	1 s	< 30 km/h	Service Level 5
Inbound logistics for manufacturing, for example, driving trajectories if supported by further sensors like camera, GNSS, Inertial Measurement Unit (IMU) of autonomous driving systems	< 30 cm (if supported by further sensors like camera, GNSS, IMU)	99.9 percent	N/A	10 ms	< 30 km/h	Service Level 6
Inbound logistics for manufacturing (for storage of goods)	< 20 cm	99 percent	N/A	< 1 s	< 30 km/h	Service Level 7

3.2.2 SECURITY REQUIREMENTS

As noted in section 2.2.12, security assurance is critical for device users and particularly so for the IoT. Security requirements are well established for the variety of application and are well-addressed by the 3GPP standards.

[134] 3GPP TS 22.104 V16.1.0, *Service requirements for cyber-physical control applications in vertical domains; Stage 1*. March 2019. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

3.2.2.1 IOT THREAT SURFACE WITH 5G

A 2017 study[135] to investigate the impact of IoT security on IT and line-of-business (LoB) leaders revealed that IT and LoB leaders are anxious about IoT security because attacks can significantly affect critical business operations. One troubling fact revealed was that when it comes to IoT, the majority of organizations cannot provide a complete accounting of all their network-connected devices even as each new device that comes online represents another expansion (another attack vector) of the overall threat surface. Even for identified IoT entities, the ownership from a security point of view frequently remains murky, further compounding the problem. At the same time 90 percent of the companies expected an increase in the volume of connected devices.

In 2016, hackers launched some of the biggest cyberattacks in internet history. These DDoS attacks were executed by infecting multiple internet-connected devices (for example, surveillance cameras, Digital Video Recorders (DVRs), routers) and then using them used to launch coordinated DDoS assaults on an array of targets, including web hosting service providers and journalists. This was named the Mirai virus. The disturbing fact about Mirai, which became clear when the source code was later revealed, was the relative lack of programming sophistication involved. Launching this botnet of things attack did not require a high degree of programming skills. The basic tools are easily available and accessible to all on the internet. The main focus of the Mirai event was that it highlighted key IoT security issues.

The four broad principles that are worthy of note for securing IoT infrastructure are:

1. Securing IoT should not be an afterthought. IoT security needs to be addressed at the design phase, not added post deployment.
2. Whether it is healthcare, automotive, energy, IoT intrinsically involves multiple layers of security: hardware, software, in-transit data, storage, network, application, and etcetera. The importance and interplay between these layers are highly contextual. Overall IoT security design must take this fact into account.
3. IoT security can only be as strong as its weakest link. Significant attention is often paid towards securing a mobile phone while ignoring what happens within the sprinkler control or car key applications that reside on it.
4. Complex IoT devices (for example, industrial equipment, connected cars) are the most difficult IoT environments to secure. Also, the consequences of hacked connected car, for example, can be substantially more serious compared to that of a connected electric meter or refrigerator.

The 5G Americas security paper explains the threat surface created by the introduction of IoT in the following sections. Comprehensive IoT security needs to consider security at many levels, as Figure 3.1 illustrates. The devices and network/transport may be the areas of primary focus today but from a revenue standpoint, the platforms, applications and services will be key. While the scope of this section is focused on IoT security in the context of 5G, it is worthwhile to take a brief look at the comprehensive IoT security landscape.[136]

[135] *IoT and OT Security Research Exposes Hidden Business Challenges,* Forrester Consulting report commissioned by Forescout Technologies, Inc. 2017. https://www.forescout.com/iot_forrester_study/
[136] *IoT security protecting the networked society*, Ericsson white paper. https://www.ericsson.com/en/white-papers/iot-security-protecting-the-networked-society

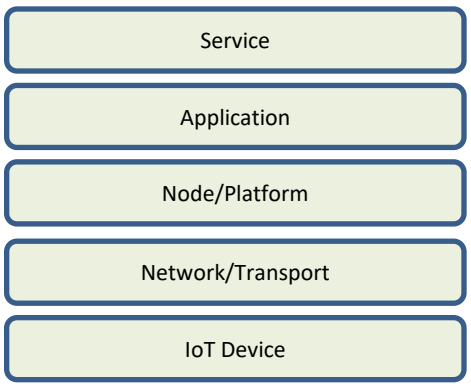


Figure 3.1. IoT Security Levels.[137]

IoT Device - Many IoT devices will likely reside in exposed and vulnerable environments. Device resident sensitive data can be tampered with. Malicious updates of device firmware and Operating Systems (OS) pose a significant problem.

Network/Transport - Network connectivity enables secure interaction of device/apps with serving network nodes. To secure this interaction, we need secure identification/authentication (credentials) and data transport. IoT network connectivity must handle billions of devices, involving heterogeneous access technologies and capillary networks, cost effectively.

Node/Platform - IoT platforms must ensure the security of data and control commands. In addition, platforms are also responsible for ensuring isolation between devices and users and third-party apps and platform-based services. Privacy concerns are one of the main inhibitors to adoption.

Application - Applications can be seen as a combination of micro services used to create a service. These applications can be statically located or dynamically migrated to the environment that is optimal for their realization. The security of the applications will be the result of the application code itself and the platform it is using. In cases where applications can migrate, it is important that migration between platforms happens securely.

Service - IoT enables a multitude of new services. A key new service in which IoT will play a significant role, and where ensuring security is of paramount importance, is connected cars. For large groups of connected vehicles traveling at high speeds, safety will always remain as a focus area. If network connectivity is lost, either because of malfunction or jamming, there needs to be backup mechanisms that on which the service can fall back. There are many other sensor-based services, of various degrees of criticality that could be enabled by IoT. The path to securing various IoT services will need to consider their uniqueness, as well as criticality of the service itself.

Evolved 3GPP technologies such as LTE-M, NB-IoT and EC-GSM-IoT are superior solutions designed to meet IoT requirements. They provide global connectivity and offer unrivaled robustness compared with unlicensed spectrum. The use of encryption on the radio interface makes traffic analysis significantly harder.[138]

[137] *IoT security protecting the networked society*, Ericsson white paper. https://www.ericsson.com/en/white-papers/iot-security-protecting-the-networked-society
[138] *Ibid.*

3.2.2.2 5G THREAT SURFACE FOR MASSIVE IOT

MIoT spans a wide variety of new and exciting opportunities, such as autonomous vehicle communications, smart grids, highway/traffic sensors, drone communications, medical sensors and AR/VR. The MIoT market opportunity, and its unique requirements and cybersecurity considerations, are directly influencing 5G architecture. Two examples are 5G's use of edge computing and its support of Ultra Reliable Low Latency Communications (URLLC).

In addition to the new opportunities and capabilities, 5G creates new cybersecurity considerations. Its use of the cloud and edge computing, and convergence of mobile and traditional IT networks, create new attack vectors. The 5G Americas whitepaper, *Wireless Technology Evolution Towards 5G*, explores how 5G provides a new set of visibility and control elements to help operators protect their networks, business partners and customers.[139]

3.3 EVOLUTION FROM 4G IOT

Many enhancements to support IoT have been proposed in the 3GPP standards from Release 12 to Release 15. While these enhancements are not considered 5G, they do form the basis on which 5G IOT builds. LTE Release 13 introduced bandwidth-reduced low-complexity (BL) and coverage-enhanced (CE) UEs to address the requirements from MTC (Machine Type Communication) and IoT applications. Release 14 introduced many further enhancements to support higher data rate, multicast, enhanced positioning and mobility as well as enhanced Voice-over-LTE (VoLTE) support. Release 15 added even more further enhancements to support new use cases, reduced latency and device power consumption, as well as improved spectral efficiency and access load control. As the number of deployed MTC networks and the volume of connected devices grow exponentially, Release 16 will add even further enhancements to support a wide range of applications, and improve the network operation and spectral efficiency in addition to the further enhancement to improve power consumption, latency, mobility, more dynamic access and scheduling.[140,141]

Internet of Things is an evolution from Machine Type Communications. MTC is basically a machine-to-machine and one-to-one communication network that allows a connected MTC device to be monitored and controlled by a centralized server without human intervention. Today's use cases usually do not require high data throughput and can tolerate high latencies, such as remote utility bill collection through smart meters or remote monitoring of patient health and fitness information. At this stage, the main functionality of the MTC network is to provide point-to-point connectivity for machines. As the new 5G era is approaching, Machine Type Communication has evolved into a multi-dimensional Internet of Things that provides ubiquitous network connections for different type of IoT devices and supports a wide range of use cases. 5G network will provide a scalable and flexible platform to support multiple services via cloud, network slicing and cognitive technologies. In addition to traditional narrow band low speed machine type services, 5G-IoT is supposed to also support URLLC (Ultra-Reliable Low Latency Communications) and mission critical services that require ultra-low latency and ultra-high reliability, such as industrial automatic control or autonomous self-driven vehicles. This is often referenced as Critical IoT. Voice over LTE-M will make it possible to support voice-enabled IoT services. Given that the number of connectivity links is expected to grow dramatically, massive IoT (M-IoT) has become one of the main focuses for 5G IoT technologies.

[139] *Wireless Technology Evolution Towards 5G*, 5G Americas Whitepaper. February 2017.
[140] 3GPP RP-181450, New WID on Rel-16 MTC-enhancements for LTE. June 2018.
[141] 3GPP RP-180371, New WI proposal: NB-IoT evolution and even further enhancements. March 2018.

3GPP standards for MTC have evolved from Release 12 to Release 16 to meet the requirements of new applications of 5G-IoT. In Release 12, device category CAT0 was defined for MTC. Compared to regular CAT 3/4 LTE devices, CAT0 has a reduced performance requirement that meets the needs of many machines while significantly reducing complexity. In addition, PSM (Power Saving Mode) and EAB (Extended Access Barring) were defined for purpose of saving battery life and preventing overload of the network.

In Release 13, enhanced Machine-Type Communications (eMTC) and Narrowband IoT (NB-IoT) were introduced, including new UE categories CAT-M1 and CAT-NB1. They are complementary narrowband IoT technologies in LTE intended for reduced complexity and cost for different types of use cases. These technologies can be deployed in the existing LTE bands with either in-band or guard band operations so that operators do not need to find new spectrum to provide MTC services. To further reduce UE power consumption, eDRX (extended Discontinuous Reception) was added, and a new power class of 20 dBm (decibels/milliwatt) was defined. Transmission Repetition and Transmit Time Interval (TTI) bundling were supported for extended coverage. CE (Coverage Enhancement) levels were defined for the MTC/IoT devices, which basically determines the number of transmission repetitions of UEs.

Based on Rel-13, more enhancements of eMTC and NB-IoT have been added into Rel-14 in order to improve mobility, positioning, spectral efficiency and support higher data rate and more delay sensitive services. New features such as Hybrid Automatic Retransmission (or Repeat) Request Acknowledgement (HARQ-ACK) bundling, larger TBS (Transport Block Size) and multiple HARQ processes were added to increase data rate and reduce latency. VoLTE over LTE-M was introduced in Rel-14 to support voice-enabled eMTC services. Techniques such as DL repetition, new repetition factors and adjusted scheduling delays were introduced to increase VoLTE coverage for eMTC devices. To support massive number of MTC/IoT devices, Release 14 uses LTE SC-PTM (Single Cell Point to Multipoint Transmission) for DL multicast transmission for eMTC and NB-IoT. In addition, a new MTC device category CAT-M2 was introduced to support IoT services that require wide bandwidth. A new power class with even lower transmit power (14 dBm) was also added for NB-IoT devices.

Release 15 standards are built upon Rel-13 and 14 features (for example, low-complexity UE categories M1 and M2, and CE Modes A and B) by adding support for new use cases and enhancements for further improvements with respect to latency, power consumption, spectral efficiency, access control and mobility as well as narrowband Reference Signal Received Power (RSRP) measurement accuracy. Specifically, use cases with much higher UE velocities (up to 240 km/h at 1 GHz and 120 km/h at 2 GHz) would be supported. To increase spectral efficiency and better support new wideband IoT applications, Rel-15 supports higher order modulation, more efficient resource allocation and reduced inter-cell interference. Latency is further reduced by reducing system acquisition time and signaling for HARQ feedback, quick release of RRC connection as well as the EDT (Early Data Transmission) feature. Further reduction of UE power consumption is achieved through new features such as:

- A new power class with a reduced transmit power of 14 dBm for CAT-M1 and CAT-M2
- Relaxed monitoring of cell reselection
- Wake-Up Signals (WUS) to reduce UE power consumption in idle mode
- EDT (Early Data Transmission)
- Reduced uplink transmission for signaling of paging and HARQ feedback

To support a wide range of IoT use cases, Rel-15 extended the NB-IoT operation to small cells, standalone operation mode and on TDD bands (such as Band 41). The mixed standalone operation allows small slices of non-LTE spectrum to be used as a standalone NB-IoT carrier and to be linked with in-band or guard NB-IoT associated with LTE spectrum. In addition, a new Physical Random-Access Channel (NPRACH) format

was introduced with a subcarrier spacing of 1.25 kHz and a cyclic prefix of 800 µs. Together with frequency hopping, cell range can be extended to up to 120 km.

As more and more new IoT applications come into the roadmap of 5G IoT, Release 16 and 17 standards will support the IoT ecosystem with the deployments and the new device types, while continuing the efforts of optimization. In addition to device power consumption and latency, new features will be defined to improve mobility, access control and scheduling for the purpose of fulfilling a wide range of requirements for different use cases and traffic models. For some use cases such as underground gas pipeline monitoring, link budget for extreme coverage of more than 20 dB and very long battery life of more than 10 years might be needed. Capabilities of connecting the RAN to the 5G core network would be necessary to provide NB-IoT services continuously within long periods of time. A well-developed 5G network would make it possible to converge the eMTC and NB-IoT technologies to a unified IoT network to support both low cost IoT devices and diversified IoT use cases. NR networks would be deployed to provide both enhanced Mobile Broadband (eMBB) and IoT services by sharing the resources dynamically.

4. 5G IOT SOLUTIONS

Beginning in Rel-16, 3GPP has taken a focus on enhancing the 5G network architecture and NR radio to better support IoT devices used by industry, enterprise, and in the home. Examples of these enhancements include introducing capabilities to support:

- ultra-reliable and low latency communications
- time sensitive communications
- interactions with Ethernet
- non-public networks
- use of licensed and unlicensed spectrum

4.1 5G ARCHITECTURE

3GPP SA2 is progressing the architecture enhancements needed to support ultra-reliable and low latency communication from the Rel-16 study phase into normative work. The following architecture discussion reflects the current views, which will continue to be developed in SA2 as the normative work for Rel-16 continues through the end of the year. Due to the heavy workload, many topics are already planned to continue towards completion in Rel-17.

4.1.1 ULTRA-RELIABLE AND LOW LATENCY COMMUNICATION (URLLC)

URLLC is critical for IoT applications not only in the Industry 4.0 use cases, but also for many of the smart city and smart home scenarios. Several enhancements have been identified to improve the performance of ultra-reliable and low latency communication (URLLC) in the 5G system. One enhancement has been addressed in Rel-16, using redundant data paths ensure communications over at least one path meet the QoS requirements. Other necessary enhancements require further study and normative work will occur in Rel-17. Three approaches to redundancy have been addressed to date: end-to-end redundant user plane paths; redundant transmissions on N3/N9 interfaces; and redundant transmission at the transport layer.

4.1.1.1 END TO END REDUNDANT USER PLANE PATHS

This subscription option allows a UE to make use of two simultaneous connections, through two next generation Node Bs (gNodeB), to two User Plan Functions (UPF), each acting as a PDU Session Anchor, as shown in Figure 4.1.

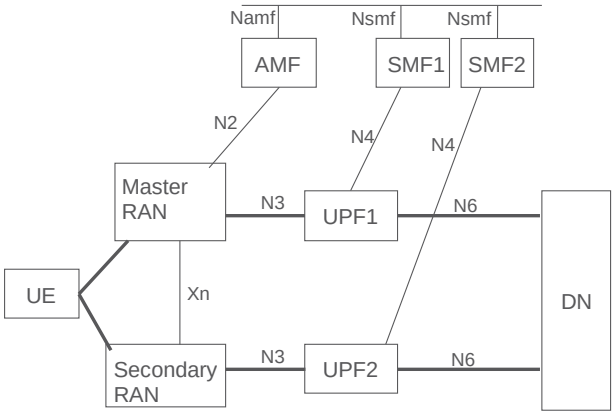


Figure 4.1. Architecture for End to End Redundant User Plane Paths Using Dual Connectivity.[142]

As the redundant PDU sessions are established, the following considerations are made:

- UE initiates the two redundant PDU sessions, providing different combinations of Data Network Name (DNN) and Single Network Slice Selection Assistance Information (S-NSSAI) for each PDU session
- The Session Management Function (SMF) determines whether the PDU sessions are to be handled redundantly. Using the combination of the S-NSSAI, DNN, user subscription and local policy configuration, the SMF determines the Redundancy Sequence Number (RSN) to differentiate the PDU sessions that are handled redundantly and determines whether the PDU session's user plane should go via the Master or the Secondary Next Generation Radio Network (NG-RAN)
- Operator configuration defines appropriate User Plane Function (UPF) selection for disjoint paths.
- The RSN parameters of each PDU session are used to request that one user plane path goes via the Master RAN, and the other user plane path goes via the Secondary RAN using dual connectivity.
- For Ethernet PDU sessions, the SMF may change the UPF, selecting a new UPF based on the identity of the Secondary RAN for the second PDU Session
- RSN information may be included in the SMF charging record
- The RSN indication is transferred from Source RAN to Target RAN in case of handover
- When a failure occurs during an attempt to establish dual connectivity the NG RAN notifies the Core Network (CN) of the failure and the SMF decides whether or not to continue with the existing PDU session[143]

[142] 3GPP TS 23.501, *System Architecture for the 5G System; Stage 2.*
[143] *Ibid.*

4.1.1.2 REDUNDANT TRANSMISSION ON N3/N9 INTERFACES

In some deployment scenarios, for example the backhaul network environment, an N3 tunnel may not be able to meet reliability requirements which are otherwise supportable in the 5G system. In such cases, the redundant transmissions between the gNB and UPF may use two independent N3 tunnels, using different transport layer paths to ensure reliability.

In such cases, the NG-RAN, SMF, or PDU Session Anchor (PSA) UPF can use different routing information to map each transport layer path to the appropriate deployment configuration, as shown in Figure 4.2.

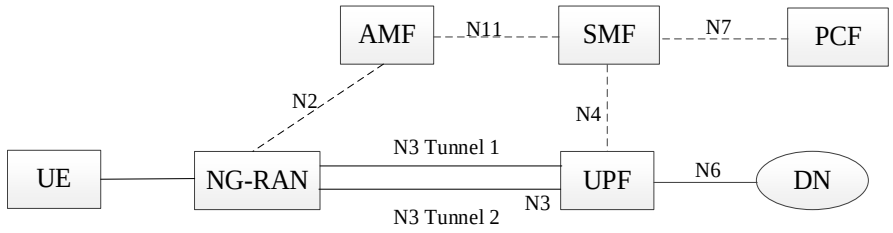


Figure 4.2. Redundant Transmission with Two N3 Tunnels Between the UPF and a Single NG-RAN Node.[144]

When this redundant N3 approach is used, the PSA UPF replicates each downlink packet of the QoS flow and assigns the same General Packet Radio System (GPRS) Tunneling Protocol User Plane (GTP-U) sequence number to each replicated packet before transmitting the packets on the individual N3 tunnels. The NG-RAN uses the sequence numbers to eliminate redundant packets before forwarding them to the UE. Similarly, uplink packets are replicated in the NG-RAN and assigned the same GTP-U sequence number before being transmitted to the PSA UPF. The PSA UPF eliminates the duplicated packets before forwarding to the DN.

The SMF may release QoS flows which require redundant transmission if a UE using dual connectivity moves to an NG RAN that cannot support redundant transmission.

An extension of this approach allows two intermediate UPFs between the PSA UPF and the NG-RAN to support the redundant transmission based on two N3 and N9 tunnels between a single NG-RAN node and the PSA UPF, see Figure 4.3. The RAN node and PSA UPF then support the packet replication and elimination function as described above.

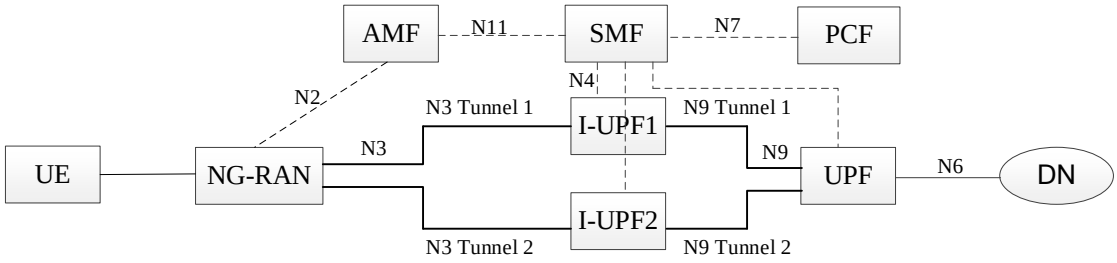


Figure 4.3. Two N3 and N9 Tunnels Between NG-RAN and UPF for Redundant Transmission.[145]

[144] 3GPP TS 23.501, *System Architecture for the 5G System; Stage 2.*
[145] *Ibid.*

4.1.1.3 REDUNDANT TRANSMISSION AT TRANSPORT LAYER

A third approach to redundancy supportable in the 5G systems uses redundant transmission at the transport layer rather than at the user plane layer or by redundant N3/N9 tunnels. An advantage of this approach is that there is no 3GPP protocol impact.

The following steps are used for redundant transport layer transmission:

- When a UE establishes the PDU session for URLLC services, the SMF selects a UPF that supports redundant functionality
- For downlink (DL) data transmission: the UPF redundancy management function duplicates the DL data on the transport layer before sends the data via the N3 GTP-U tunnel. The NG-RAN redundancy management function eliminates the duplicated DL data before forwarding to the NG RAN
- For uplink (UL) data transmission: the NG-RAN redundancy management function duplicates the UL data on the backhaul transport layer before sending the data via the N3 GTP-U tunnel. The UPF redundancy management function eliminates the duplicated UL data and forwards to UPF.[146]

4.1.2 INTEGRATION OF 5G WITH IEEE TIME SENSITIVE NETWORKING (TSN) TECHNOLOGY

Time Sensitive Networking (TSN) technology was created in the Institute of Electrical and Electronic Engineers (IEEE) to enable Ethernet links to serve a variety of demanding applications with deterministic service requirements. The applications include audio-video bridging for studios, power-line control systems, remote radio-head connectivity for wireless infrastructure and factory automation. The TSN specifications are maintained by the TSN task group under the IEEE 802.1 working group.[147]

Some of the key specifications with TSN include IEEE 802.1AS, “Timing and Synchronization” and IEEE 802.1Qbv, “Enhancements for Scheduled Traffic.” These TSN components allow multiple nodes in a TSN system to operate isochronously (therefore, operations at different nodes can be timed to occur at the same time with microsecond tolerance), and deterministically (therefore, there is a tight delay guarantee for communication among the nodes).

Within the context of the present whitepaper for IIoT, the power-line control and factory automation use-cases are relevant. Factory automation has traditionally used a variety of customized Ethernet based technologies that are not mutually interoperable. TSN technology aims to create a high-performance Ethernet based interoperable solution. Several industry organizations are working to define a TSN based interoperable solution for factory automation over Ethernet links.

- Avnu Alliance has provided a broad framework[148]
- IIC has developed a testbed for interoperability testing[149]
- OPC UA has started work for Field Level Communication profile with TSN[150]

[146] 3GPP TS 23.501, *System Architecture for the 5G System; Stage 2.*
[147] Time-Sensitive Networking Task Group, IEEE. http://www.ieee802.org/1/pages/tsn.html
[148] *Theory of Operation for TSN-enabled Industrial Systems*, AVNU Alliance. 2017. http://avnu.org/knowledgebase/theory-of-operation/
[149] *IIC Time-Sensitive Networking Testbed Continues Successes and Brings Live Demonstrator to Hannover Messe*, https://www.iiconsortium.org/press-room/03-31-19.htm
[150] *OPC Foundation Extends OPC UA Including TSN Down to the Field,* https://opcconnect.opcfoundation.org/2018/12/opc-foundation-extends-opc-ua-tsn-down-to-the-field/

4.1.2.1 TIME SENSITIVE NETWORKING OVER 5G

The benefits of wireless connectivity for IIoT have been explored in 3GPP, and are summarized in section 2.5. IIoT applications can benefit by having TSN-like performance without being tethered to wired Ethernet, enabling increased flexibility for the factory.

In 3GPP Rel-16, there is ongoing work to provide TSN performance over the 5G wireless links. This work includes the key components in Table 4.1, spread across SA and RAN working groups.

Table 4.1. 3GPP Rel-16 Work for TSN Integration with 5G.[151]

Feature	3GPP Requirement, TS 22.104	3GPP Technical work in Rel-16
Ethernet bridge architecture	Support of Ethernet PDU session type, TS 22.104 Section 6.2	Basic support in Rel-16. Enhancements in Vertical_LAN (SA)
Time Synchronization	Support of 802.1AS, TS 22.104 Section 5.6	Vertical_LAN (SA), IIoT (RAN)
Scheduling enhancements	Support of 802.1Qbv, TS 22.104 Section 6.2	Vertical_LAN (SA), IIoT (RAN)
Latency and Reliability	TS 22.104, Tables 5.2-1, 5.3-1	eURLLC (RAN)

Based on the solutions being developed in 3GPP, the 5G Alliance for Connected Industries and Automation (5G-ACIA) is working on technology requirements and regulations as well as certification for TSN uses with 5G.[152] The 5G-ACIA brings together different stakeholders for IIoT, including Operational Technology (therefore, verticals), Information Technology (therefore, telecommunications vendors) and Mobile Network Operators (therefore, service providers).

Further details of the technology components enabling TSN integration with 5G are described in subsequent sections.

4.1.2.2 MODELING OF 5G AS AN ETHERNET BRIDGE

IEEE TSN consists of a set of Ethernet bridges connected via Ethernet links. In contrast, the 5G system consists of a more diverse variety of nodes, including the gNB, UPF and UE, and a more diverse variety of links. Hence, there is a need for an architecture model that allows 5G to integrate into the TSN framework, preferably without significant changes to the nodes and interfaces that are part of 5G technology.

The architecture in Rel-16 models the entire 5G system, including the core network, the radio access network and the UE as a single Ethernet bridge. This architecture is also referred to as the “block box” option (Figure 4.4).

[151] 3GPP TS 22.104 V16.1.0, *Service requirements for cyber-physical control applications in vertical domains*; Stage 1, March 2019. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.
[152] *5G for Connected Industries and Automation Whitepaper*, 5G Alliance for Connected Industries and Automation (5G-ACIA), April 2018. https://www.5g-acia.org/publications/5g-for-connected-industries-and-automation-white-paper/

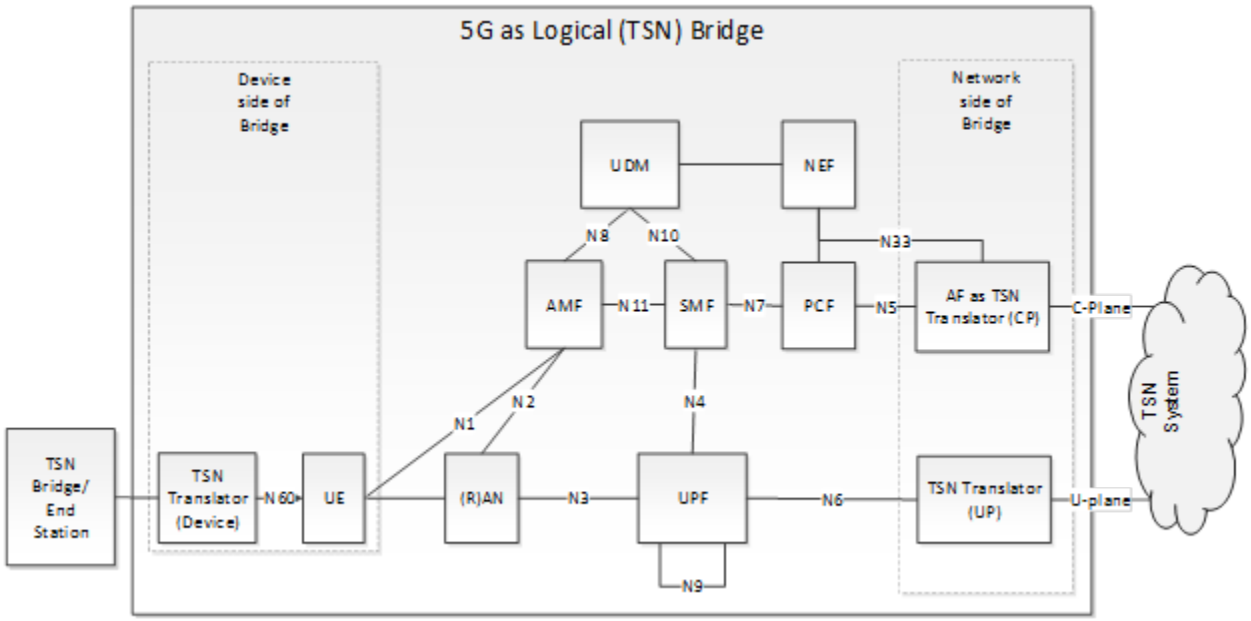


Figure 4.4. Integration of the 5G System as a Bridge.[153]

5G integration with TSN includes newly-introduced TSN translators on the device and network sides, which allow the 5G system to present itself as a bridge to TSN nodes. The advantage of this architecture option is that the TSN system does not need to be modified to interwork with 5G, allowing the deployment of 5G in industrial environments where TSN systems already exist. Such scenarios are commonly referred to as brownfield deployments in the industrial community.

The translator functions perform the following critical tasks:

- QoS management: TSN configures QoS using bridge management interfaces between a bridge and a TSN configuration system. The TSN translator on the network side converts the bridge configuration received from the TSN system into 5G-internal QoS procedures that achieve equivalent behavior in the 5G system. For example, the 5G system may create a 5G QoS flow upon receiving to a certain bridge configuration
- Dejitter at egress: TSN requires packets to be delivered from the bridge to the peer bridge at predictable time-instant. However, a wireless 5G system may sometimes deliver the packet earlier than the delivery deadline, for example, due to a good channel realization that resulted in packet delivery without use of HARQ retransmission. In such cases, the TSN translator buffers the packet until the scheduled delivery time

A more complete list of the translator functions, including Ethernet header handling can be found in 3GPP TS 23.501.

[153] From 3GPP TR 22.734, Section 6.8.

4.1.2.3 5G INTEGRATION WITH GPTP FOR TIME DELIVERY TO UE

One of the key features of TSN is the delivery of precise timing information to the UE. Industrial and powergrid use-cases require all devices as well as network entities to share a common notion of time, with uncertainty as strict as 1 microsecond. Precise timing allows different devices to execute commands and perform measurements at precisely the same time instant.

It should be noted that the delivery of time information is independent from the delivery of application packets itself. The command message itself does not need to be delivered at a microsecond precise time; rather the message has to be delivered sometime before the time instant of command execution and then the device uses its knowledge of time to execute the command at the precise time.

TSN uses IEEE 802.1AS generic Precision Time Protocol (gPTP) to deliver timing information across a TSN network. The gPTP framework involves the flow of gPTP packets across Ethernet bridges, with each Ethernet bridge adjusting a correction field in the message to account for its residence time. The final receiver of the message can derive the total time elapsed since the message was generated, and thus determine the TSN time.

Rel-16 includes functions for 5G in its role as an Ethernet bridge to be able to add residence time corrections to gPTP messages as they flow through.

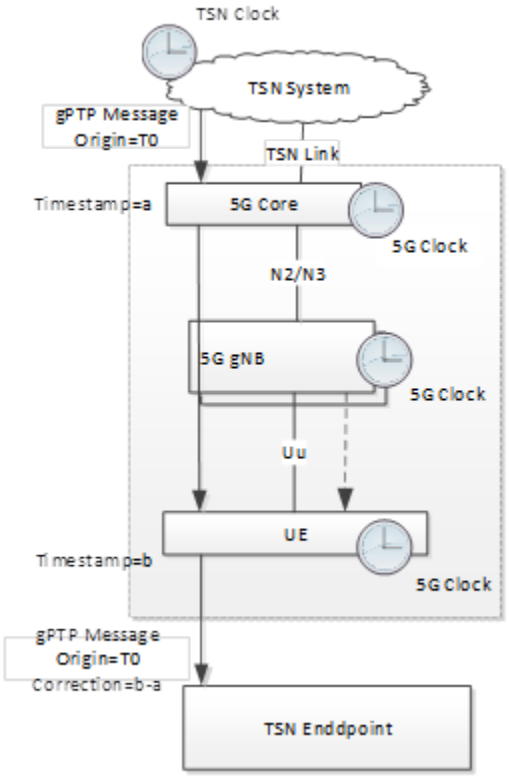


Figure 4.5. gPTP Message Delivery Via 5G.

The time delivery function in 5G relies on the UE's precise synchronization with the gNB that is part of basic air interface operation. The following steps are involved:

1. The UE receives the 5G time via air interface signaling from the gNB, and it is assumed that the 5G Core has access to the same 5G time
2. 5G core records the time the gPTP message was received at the core (say timestamp=a) and sends this value to the UE
3. The UE records the time when the gPTP message was delivered to the TSN endpoint (say timestamp=b)
4. The UE computes a correction field (b-a) and delivers it to the TSN endpoint
5. The TSN endpoint can adjust the received origin time in the gPTP message by the correction field to derive an accurate TSN time

The timing delivery function follows the overall "black box" architecture approach for TSN over 5G, and does not require the TSN system or TSN endpoints to deviate from standard IEEE 802.1AS procedures.

4.1.3 FLEXIBLE DEPLOYMENT MODELS FOR NON-PUBLIC NETWORKS

Service requirements for Industrial IoT applications have been studied in 3GPP TR 22.804. Section 5.3.1.2 points out that IIoT applications in factories often require a high level of security and availability, and are subject to stringent liability and business constraints. This means that typical public networks may sometimes not be able to serve the needs of IIoT applications in factories.

3GPP Rel-16 defines non-public networks that can be deployed to meet the specific needs of Industrial IoT. Though the standards' defined term is 'non-public network', the terminology 'private network' is sometimes used in the industry to refer to equivalent deployment models.

The Rel-16 non-public network architecture is independent of the type of spectrum used. Spectrum options available for IIoT networks include dedicated IoT bands being allocated in some regions (for example, Germany and Sweden), shared bands (for example, CBRS for the U.S.) and unlicensed bands (for example, 5 GHz for NR-unlicensed operation). In addition, spectrum holders can allocate some of their spectrum in specific areas for non-public networks.

These non-public networks can be standalone and independent of any external networks or could be non-standalone with a relationship with an external public network. In the case of a non-standalone non-public network, it is also possible to provide mobility to devices that move in and out of the factory, for example, for supply chain management scenarios.

A large set of identifiers is available for non-public networks, which allows multiple such networks to be deployed. UEs can be provisioned with network-specific credentials that restrict the UE to only access the particular non-public network that it has a subscription for. If a non-public network sees an access attempt from a UE that does not have the correct subscription, the non-public network will reject the UE.

4.1.4 3GPP IOT PROTOCOLS

Support for IoT was introduced in 3GPP specifications as part of eMTC and NB-IoT features in Rel-13. These features focused on enhanced coverage, battery life and scalability, and included radio as well as core network functions.[154]

With the introduction of 5G in Rel-15, a new core network (5GC) was defined. The benefits of 5GC include a more "clean slate" approach to incorporate technology trends such as virtualization, control and user

[154] *LTE Progress Leading to the 5G Massive Internet of Things*, 5G Americas Whitepaper, December 2017.

plane separation, edge computing, and as summarized below from Section 5.3 of a recent 5G Americas whitepaper:[155]

> While revolutionary in most respects, in other respects the 5G architecture is evolutionary, building on trends already in flight, for example, Control and User Plane Separation (CUPS), network function virtualization, and etcetera. Rel-14 enabled new capabilities in a system built largely on the network architecture and elements defined in prior releases. The 5G Next Generation Core (NGC), beginning with Rel-15, takes a more "clean slate" approach, designing in the concepts of CUPS, Network Slicing, Edge Computing and virtualization from the start, and also defines an elegant and flexible methodology for incorporating non-3GPP access mechanisms.

In Rel-16, the eMTC and NB-IoT radios can be connected to the 5GC, allowing service providers increased flexibility of deployment as the 5G deployments progress.

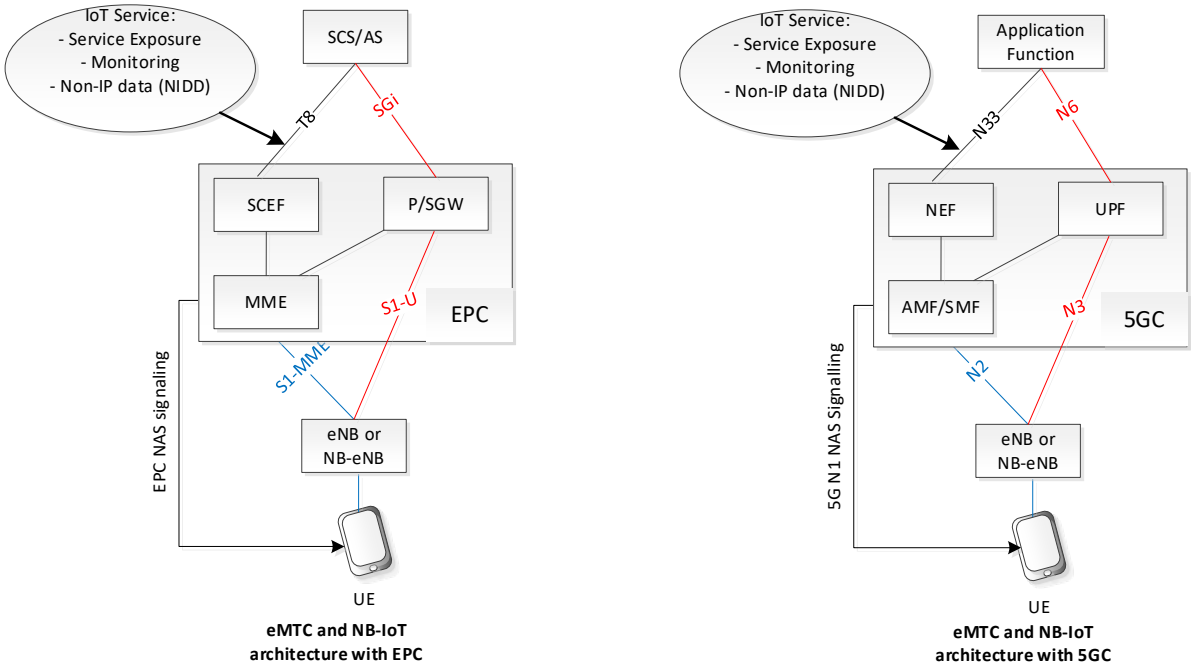


Figure 4.6. IoT Architectures with EPC and 5GC.

As described in Section 4.1.5 of a previous 5G Americas whitepaper,[156] the key features for CIoT available with the EPC are:

- Dedicated Core Networks (DECOR)
- Architecture Enhancements for Services capability exposure (AESE)
- Optimization to support High Latency Communication (HLCom)
- Group Based Enhancements (GROUPE)
- Monitoring Enhancements (MONTE)
- Architecture Enhancements for Cellular Internet of Things (CIoT)

[155] *Wireless Technology Evolution: Transition from 4G to 5G,* 5G Americas Whitepaper. October 2018.
[156] *LTE Progress Leading to the 5G Massive Internet of Things*, 5G Americas whitepaper. December 2017.

3GPP Rel-16 makes these features available with 5GC also. Details are available in Section 5.31 of 3GPP TS 23.501.

4.2 SPECTRUM OPTIONS FOR LICENSED, UNLICENSED AND SHARED BANDS

Over the years, consumers have been quick to adopt wireless technologies made possible by airwave frequencies that deliver everything from radio and TV broadcasts, to Wi-Fi Internet connections and smart phones. Yet, as more people become armed with more wireless devices, the congestion across these frequencies increases and the performance of the signals begins to erode.

Spectrum is divided into the two main classes of licensed and unlicensed. The flood of new IoT devices and applications coming onto the market are made possible by the latest advances in connectivity, both in licensed and unlicensed spectrum. Unlicensed technologies include short-range technologies like ZigBee and long-range technologies like LoRa. Licensed technologies include Narrowband IoT (NB-IoT), LTE for Machine Type Communications (LTE-M) and Enhanced Coverage GSM (EC-GSM).

With so many licensed and unlicensed connectivity choices, it's natural to consider whether there is a need for all of them. Certainly, when you look at the technical specifications for Low-Power Wide Area (LPWA) technologies, the differences are quite small. But looking only at technical specifications would be missing the point of IoT. It's not about the technology, it's about the application that enhances our lives. Licensed and unlicensed connectivity options are, therefore, truly complementary and it is believed that both are needed if the Internet of Things is to live up to its full potential.

4.2.1 LICENSED SPECTRUM

Mobile operators already provide reliable, end-to-end secured IoT platforms that allow customers to scale and manage their business requirements. They also have unrivalled global network coverage as well as technical and business support to react to a customer's changing needs. As trusted providers of reliable solutions, operators and their ecosystem partners are therefore best placed to extend their reach to serve the full range of IoT applications. This means solutions deployed in licensed spectrum will:

- Support very low in power consumption – a battery life measured in years, in excess of 10 years for some applications
- Be optimized for brief messages – about the length of an SMS
- Have a very low device unit cost – the connectivity module will eventually cost a few dollars according to many industry analysts
- Have good coverage both indoors and outdoors and in previously unreachable locations, often beyond power sources
- Be easy to install on to current networks, reusing existing cellular infrastructure wherever possible
- Be scalable by being able to support large numbers of devices over a wide geographic area
- Deliver end-to-end secure connectivity and support for authentication appropriate to the IoT application
- Be able to be integrated into a mobile operator's unified IoT platform

There is a good, and growing amount of licensed mobile coverage (therefore, sub-1 GHz) and capacity (therefore, above 1 GHz) spectrum to support the rapid growth of IoT. Mobile services in these bands are well established worldwide in mature networks and can be employed to support IoT as well as personal mobile services relatively easily. Licensed spectrum includes: Low band (below 1 GHz); mid band (sub 6 GHz); and high band (higher than 6 GHz). In practice, most of the bands that will be used for cellular IoT

will more likely be sub 1 GHz for IoT applications requiring wide range. Table 4.2 shows a list of licensed low bands commercially deployed around the world.

Table 4.2. List of Licensed Low Bands.

3GPP band #	Frequency (MHz)		Max BW	Region	Duplexing
	UL	DL			
5	824-849	869-894	10	North/Latin America, japan, Australia	FDD
26	817-824	862-869	5	U.S- Sprint	FDD
8	880-915	925-960	10	Global (excluding North America)	FDD
12	699-716	729-746	10	North/Latin America	FDD
13	777-787	746-756	10	North/Latin America	FDD
14	788-798	758-768	10	North/Latin America	FDD
17	704-716	734-746	10	North/Latin America	FDD
71	663-698	617-652	20	U.S	FDD

Table 4.3 depicts a list of licensed mid bands, which includes bands for Frequency Division Duplex (FDD) as well as Time Division Duplex (TDD) operation.

Table 4.3. List of Licensed Mid Bands.

3GPP/Band		Frequency (MHz)		Max BW	Region	Duplexing
		UL	DL			
2	PCS	1850-1910	1930-1990	20	North, Latin America	FDD
25	PCS	1850-1915	1930-1995	20	U.S- Sprint	FDD
3	DCS	1710-1785	1805-1880	20	Global (excluding North America)	FDD
4	AWS-1	1710-1755	2110-2155	20	North, Latin America	FDD
66	AWS-1-3	1710-1780	2110-2200	20	U.S.	FDD
70	AWS-4	1694-1710	1995-2020	15	U.S.	FDD
41	BRS	2496-2690	2496-2690	20	U.S. (Sprint), China, japan	TDD
42	CBRS	3400-3600	3400-3600	20	Europe, Japan	TDD
48	CBRS	3550-3700	3550-3700	20	U.S.	TDD
49	CBRS	3550-3700	3550-3700	20		TDD

There has been significant progress around the world to secure millimeter Wave (mmW) for 5G networks. In the U.S., significant progress has been made toward making mmW spectrum available for 5G. These bands have traditionally been used for fixed and satellite services. The FCC has been driving this process in several steps.

- Notice of Inquiry issued end of 2014
- Notice of Proposed Rulemaking (NPRM)issued end of 2015
- Report &Order (R&O) and further NPRM, issued July 14, 2016
- Second R&O, and second NPRM, issued November 22, 2017

mmWave is one of the key enablers for 5G systems, and appears to be a promising candidate for next generation wireless networks delivering multi gigabit-per-second data rates.[157] [158] A challenge facing mmWave is due to its high carrier frequency resulting in higher propagation loss. Factors such as rain attenuation, atmospheric absorption and wall penetration losses can limit the range of mmWave. Free space propagation loss is proportional the square of the frequency carrier, in addition, non-line-of-sight (NLOS) suffers higher attenuation than the line-of-sight (LOS) system.[159] This results in smaller cell sizes on the order of 200 m^2, which might hinder its application for some services. However, smaller cell sizes would provide higher spectral efficiency and lower latency. mmWave could be a good fit for use cases requiring localized coverage, e.g. smart cities, enterprise and industrial IoT applications.[160]

More information on licensed spectrum is available in 5G Americas white paper, *5G Spectrum Vision*, published in February 2019.[161]

4.2.2 UNLICENSED SPECTRUM

A primary technology deployment in unlicensed spectrum is WiFi. WiFi provides small coverage areas resulting in high-frequency reuse and high data density (bps per square meter). Less efficient are white-space unlicensed networks, sometimes called "super Wi-Fi," that, because of large coverage areas, have much lower throughput per square meter. While white-space networks may be a practical broadband solution in rural or undeveloped areas, they face significant challenges in urban areas that require higher throughput density and already have mobile and fixed broadband available.

Zigbee is another local area unlicensed technology as well as Bluetooth Low-Energy which is more usable for personal area coverage. Low-Power Wide-Area (LPWA) technologies emerging specifically for the IoT market are Ingenu, LoRa and Sigfox. Table 4.4 summarizes the LPWA technologies, WiFi and cellular technologies for IoT based on the whitepaper by Rysavy Research published in October 2018.

Table 4.4. Wireless Networks for IoT.[162]

Technology	Coverage	Characteristics	Standardization/ Specifications
GSM/GPRS/EC-GSM-IoT	Wide area. Huge global coverage.	Lowest-cost cellular modems, risk of network sunsets. Low-throughput.	3GPP
HSPA	Wide area. Huge global coverage.	Low-cost cellular modems. Higher power, high throughput.	3GPP

[157] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, et al., *Millimeter wave mobile communications for 5G cellular: It will work!,* IEEE access, vol. 1, pp. 335-349, 2013.
[158] E. Hossain and M. Hasan, *5G cellular: key enabling technologies and research challenges*, IEEE Instrumentation & Measurement Magazine, vol. 18, pp. 11-21, 2015.
[159] E. Hossain and M. Hasan, *5G cellular: key enabling technologies and research challenges*, IEEE Instrumentation & Measurement Magazine, vol. 18, pp. 11-21, 2015.
[160] *Spectrum for the Internet of Things*, GSMA public policy position, August 2016 and *Mobile IoT*, GSMA are used in general reference in this subsection on Licensed Spectrum.
[161] *5G Spectrum Vision*, whitepaper by 5G Americas. February 2019.
[162] *LTE to 5G – The Global Impact of Wireless Innovation,* Rysavy Research. October 2018.

Technology	Coverage	Characteristics	Standardization/ Specifications
LTE, NB-IoT	Wide area. Increasing global coverage.	Wide area, expanding coverage, cost/power reductions in successive 3GPP releases. Low to high throughput options.	3GPP
Wi-Fi	Local area.	High throughput, higher power.	IEEE
ZigBee	Local area.	Low throughput, low power.	IEEE
Bluetooth Low Energy	Personal area.	Low throughput, low power.	Bluetooth Special Interest Group
LoRa	Wide area. Emerging deployments.	Low throughput, low power. Unlicensed bands (sub 1 GHz, such as 900 MHz in the U.S.)	LoRa Alliance[163]
Sigfox	Wide area. Emerging deployments.	Low throughput, low power. Unlicensed bands (sub 1 GHz such as 900 MHz in the U.S.)	Sigfox[164]

3GPP is working to develop forms of 5G NR designed from the ground-up to operate in unlicensed and shared spectrum. Thus, NR-U will be able to operate in existing unlicensed bands or in green field unlicensed or shared bands. Such applications typically require low power platforms of low cost, but are able to transmit messages at reasonable distances.[165]

4.2.3 SHARED SPECTRUM

Spectrum sharing is the simultaneous usage of a specific radio frequency band in a specific geographical area by a number of independent entities. Simply, it is the "cooperative use of common spectrum" by multiple users. Spectrum sharing matters because communications spectrum is a scarce asset, and demand is growing, due to user growth, new apps and devices consuming more bandwidth. Spectrum sharing also can take many forms, coordinated and uncoordinated.

The 3GPP study provides a great opportunity also to explore new sharing paradigms targeting green field shared/unlicensed spectrum that can deliver significant benefits in terms of increased spectral efficiencies,

[163] For details, see LoRa Alliance, https://www.lora-alliance.org/.
[164] For details, see Sigfox, https://www.sigfox.com/en.
[165] *Wireless Innovation-From LTE-U/LAA to 5G spectrum sharing,* Qualcomm. March 2018 and *Long range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios,* Marco Centenaro, Lorenzo Vangelista, Andre Zanella, Michele Zorzi, IEEE wireless communication. October 2015 are used in general reference in this subsection on Unlicensed Spectrum.

higher perceived user data speeds, and guaranteed bandwidth and Quality of Service (QoS) than is possible today. Coordinated forms include:

- Capacity sharing between business entities (roaming, wholesale, pooling of assets)
- TV white spaces (database determines what you may use, when and where)
- Spatial sharing between business entities
- Priority sharing between entities
- LAA (bonding of mobile and Wi-Fi assets)
- Cognitive radio (devices determine how to avoid interference)

In some cases, sharing is a business arrangement between entities. Historically, Mobile Virtual Network Operator (MVNO) wholesale is a form of sharing. So too is "roaming" in a sense. In other cases, mobile operators might agree to pool and share licensed spectrum assets. The arguably more important forms of spectrum sharing use new technology to intensify the use of existing spectrum that allows many users to share a specific block of spectrum. The concept is to free up capacity quickly by allowing commercial users access to currently-licensed spectrum on a secondary basis, while licensed users continue to retain priority use of their spectrum.

Such sharing allows licensed services to share spectrum in a band with new users without disrupting existing users, while still increasing the amount of spectrum available for other users. Policy makers are evaluating how spectrum might be shared between government and commercial entities.

As an example, the U.S. government can designate spectrum for exclusive, shared, or unlicensed use, as shown in Figure 4.7. Shared use can be opportunistic, as with TV white spaces; two-tier with incumbents and licensed users; or three-tier, which adds opportunistic access. The bands initially targeted for spectrum sharing in the U.S. include AWS-3 (two tiers on a temporary basis) and the 3.5 GHz band (three tiers).

The three-tier plan envisioned by the U.S. government for the 3.5 GHz band gives more entities access to the spectrum but at the cost of increased complexity.

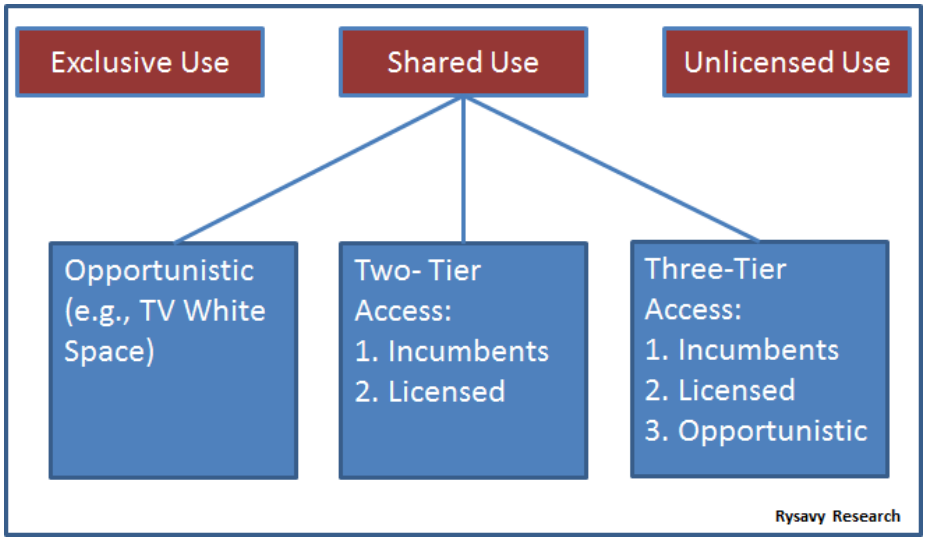


Figure 4.7. Spectrum Use and Sharing Approaches.[166]

Although a potentially promising approach for the long term, sharing raises complex issues. For example, the United States is developing an approach to sharing in the 3.5 GHz band. In the U.S. model, a three-

[166] TV White Space are under FCC Unlicensed Part 15 rules, Subpart H.

layer model is envisioned, with protected incumbent access, priority access (some interference protection) and general authorized access (opportunistic access without interference protection).

Spectrum sharing may be required in some 5G bands, including 38.6 to 40 GHz, such as with fixed satellite service.

4.2.4 CHOICE OF FREQUENCY BANDS FOR IOT APPLICATIONS

In addition to supporting traditional services provided by the existing mobile networks, three new service categories are envisioned for 5G – enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low-Latency Communications (URLLC) and massive Machine-Type Communications (mMTC). These three new categories of services have diverse requirements in terms of bandwidth, latency, mobility, connection density and data rates. For example, eMBB places high requirements on spectrum efficiency, peak data rate, area traffic capacity and network energy efficiency.

All frequency bands (low, mid, high), whether licensed or unlicensed, will play a role in future 5G networks. Applications requiring wide range and mobility will be best served by low/mid band licensed spectrum, where we already have mature networks in place. High band will play a role for applications requiring higher throughput and lower end-to-end latency, and improve coverage. Table 4.5 provides a view of different category of performance requirements.

Table 4.5. Different Levels of Performance Requirements.[167]

	Low	Mid	High
End-End latency	> 1 sec	10-100 msec	<1 msec
Data rate- Mbps	< 1 DL, 256 Kb UL	10 DL/1 UL	>100 DL, 10 UL
UE speed- Km/h	< 3	<75	>75
Reliability	<99 percent	99.9-99.99 percent	>99.999 percent
Cell range	<200m	500m-1Km	>1Km

Table 4.6 presents a table of performance level with frequency band of operation.

Table 4.6. Performance Level vs. Frequency Bands.

Performance level	Frequency band	Low	Mid	High
Cell range	High	x	x	
	Mid	x	x	
	Low	x	x	x
Data rate	High		x	x
	Mid	x	x	x
	Low	x	x	x
Latency	High			x
	Mid	x	x	x
	Low	x	x	x
UE speed	High	x	x	

[167] 3GPP TS22.186 V16.1.0, *Enhancement of 3GPP support for V2X scenarios* and 3GPP TS22.104 V16.1.0, *Service requirements for cyber-physical control applications in vertical domains; Stage 1*. These requirements for design targets may not be met by the Rel-16 specifications scheduled for completion by end of 2019.

	Mid	x	x	
	Low	x	x	x
Reliability	High	x	x	
	Mid	x	x	x
	Low	x	x	x

4.3 NEW RADIO ENHANCEMENTS FOR ULTRA-RELIABLE AND LOW LATENCY COMMUNICATIONS

Based on the 3GPP requirements presented in Section 3, lots of IoT use cases, such as communication with vehicles, industrial automation and process control, have very stringent requirements on system reliability and end-to-end latency. For example, motion control requires that the maximum allowable end-to-end latency needs to be 0.5 - 2 ms and communication service availability needs to be 99.999 percent - 99.99999 percent. To support this type of URLLC (Ultra-Reliable and Low Latency Communications) services, or Critical Communications, 3GPP Releases 15 and 16 have defined solutions for NR at both the physical layer and upper layers to reduce latency and improve system reliability. This section will illustrate some of the key URLLC techniques for 5G NR.

As a very important KPI for 5G URLLC service, the end-to-end latency consists of different components - including gNB processing time, UE processing time, minimum Transmission Time Interval (TTI), time required for DL ACK/NACK transmission, PDCCH transmission latency and DL/UL data transmission for TDD blocks within the subframe. The 5G system needs to optimize data transmission procedure and minimize the latency on both control and user plane to meet the tight latency requirements.

4.3.1 FLEXIBLE NR FRAMEWORK

To simultaneously support different services with different QoS and performance requirements, 5G NR provides a flexible frame structure. Different numerologies with different SCS (Sub-Carrier Spacing) can be used to provide multiple services with different QoS requirements. Table 4.7 shows different numerologies defined by 3GPP.[11] A higher numerology index corresponds to a higher SCS and shorter slot length, therefore it can be used to support URLLC services. Data transmission will be scheduled with much shorter slots so that transmission latency can be greatly reduced.

Table 4.7. Different Numerologies for NR.[168]

Numerology	SCS (kHz)	Slots per Subframe	Slot Length
0	15	1	1 ms
1	30	2	0.5 ms
2	60	4	250 µs
3	120	8	125 µs
4	240	16	62.5 µs

[168] 3GPP TS 38.211 V15.5.0, *NR Physical Channels and Modulation (Rel. 15).* March 2019

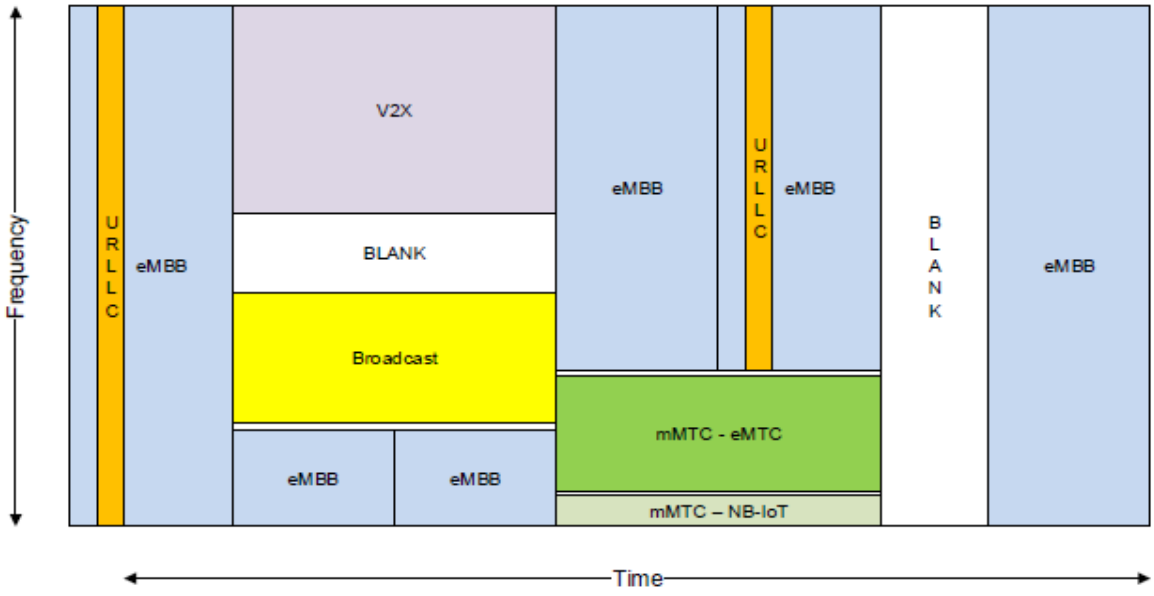


Figure 4.8. Flexible Frame Structure Of 5G NR to Support Multiple Services with Different Qos Requirements.[169]

4.3.2 NON-SLOT BASED SCHEDULING (MINI-SLOT SCHEDULING)

If small SCS (for example 15 kHz) is implemented in 5G NR, there is another option for reducing TTI to support URLLC services, which is non-slot-based (or mini-slot based) scheduling. To reduce processing time and latency, the URLLC data is scheduled within shorter version of slots, i.e. mini-slots. Based on 3GPP definition, a mini-slot is much shorter than the standard slot size. Each mini-slot can be defined as 7, 4, or even 2 Orthogonal Frequency Division Duplexing (OFDM) symbols per slot, while a standard slot contains 14 OFDM symbols). A mini-slot has the same control channel structure as the standard slot. When there are 2, 4, or 7 symbols per slot, data transmission can start immediately without needing or waiting for standard slot boundaries, which reduces processing time and enables much faster delivery of data (see Figure 4.9). Figure 4.10 shows the procedure of scheduling based on mini-slot.

[169]Nokia Bell Lab, *Ultra Reliable Low Latency Communication for 5G New Radio*, IEEE Workshop on 5G Technologies for Tactical and First Responder Networks. Oct. 2018.

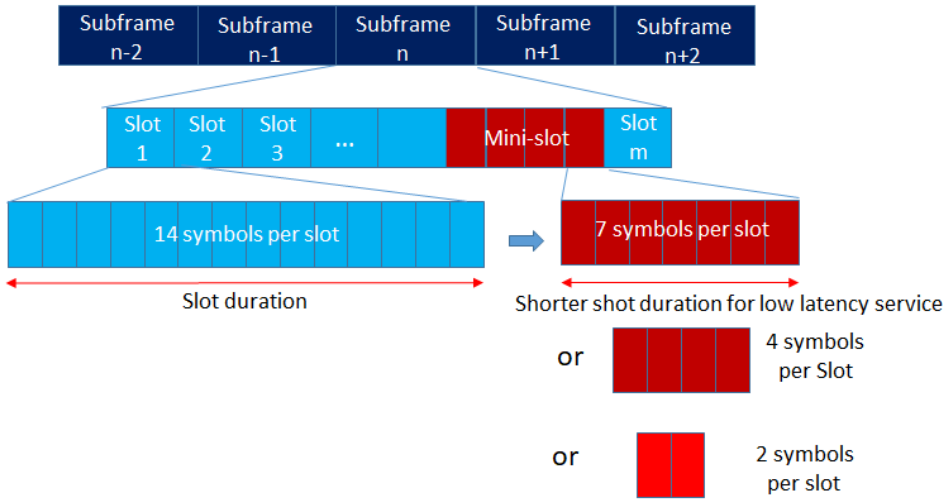


Figure 4.9. Mini-Slots in a Flexible Frame Structure in 5G NR.

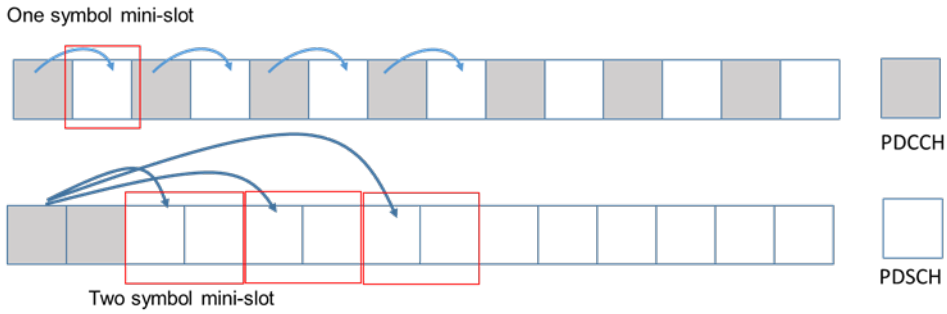


Figure 4.10. Mini-slot (non-slot) Based DL Scheduling.

To enable the performing channel estimation earlier, the front-loaded DMRS (Demodulation Reference Signal) is supported in NR. The Demodulation Reference Signal is located just after the control region and followed by data region, as seen Figure 4.11. Once the channel estimates are obtained from the front-loaded DMRS, the receiver can demodulate data in the data region right away. This will greatly reduce UE processing time and thus reduce the latency at user plane.

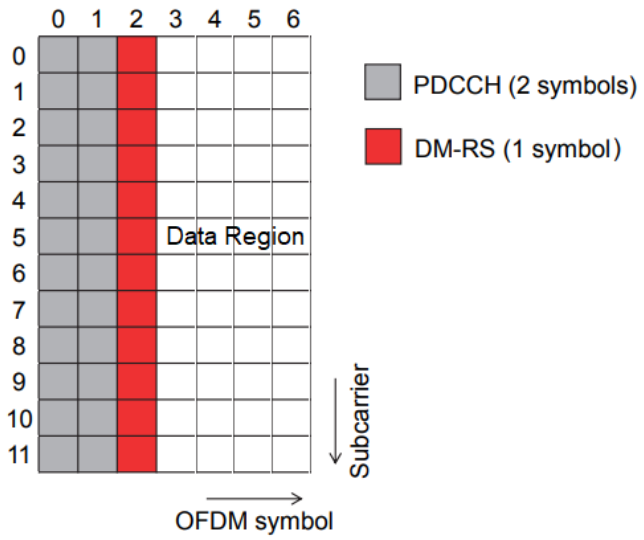


Figure 4.11. Location of Front-loaded DMRS in a Mini-slot of 7 Symbols.

4.3.3 SEMI-PERSISTANT SCHEDULING FOR DL TRANSMISSION

To reduce the scheduling assignment overhead in the control channel and reduce latency, Semi-Persistent Scheduling (SPS) is configured by RRC per serving cell and per bandwidth part. Once DL SPS is activated through L1 signaling, a DL assignment is provided by PDCCH and stored as configured grant for the serving cell. When SPS is configured, RRC configures the following parameters:

- CS-RNTI (Configured Scheduling Radio Network Temporary Identifier) for activation, deactivation and retransmission
- Number of configured HARQ processes for SPS
- Periodicity of configured DL assignment for SPS

After a DL resource grant is assigned, the Media Access Control (MAC) layer assumes this DL assignment will occur periodically until SPS is deactivated. When SPS is deactivated, all the corresponding configurations will be released. As shown Figure 4.12, SPS allows gNB to transmit DL data directly to the UE without waiting for UE request and sending grants. Compared to regular dynamic scheduling, SPS greatly reduces the additional overheads of “UE grant requests” and “DL transmission grants” so that DL packets can reach the UE faster.

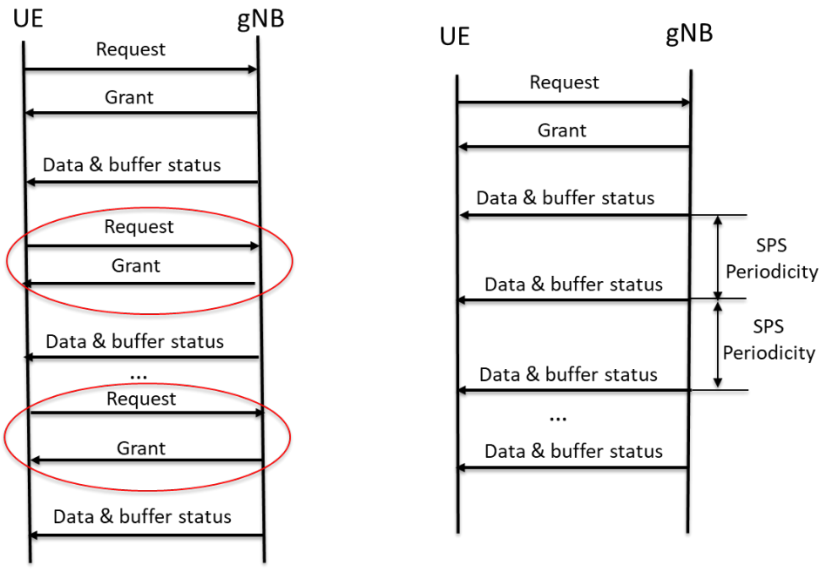


Figure 4.12. Dynamic Scheduling vs. SPS (Semi Persistent Scheduling).

4.3.4 UL GRANT FREE TRANSMISSION

The mechanism of UL grant free transmission is similar to DL SPS, except that it is used for uplink data transmission. In regular UL transmission, UE has to do the regular scheduling hand-shake. It needs to send the scheduling request first then wait for the scheduling grant before it can actually start the data transmission. To avoid this additional delay, 3GPP Rel-15 provides two types of grant-free (GF) configuration schemes, which are Configured Grant Type 1 and Type 2.[170] This procedure allows UE to perform uplink data transmission periodically without waiting for the UL grant.

In case of Configured Grant Type 1, an uplink grant is provided by RRC and stored as configured uplink grant for the indicated service cell. During RRC configuration, the following parameters are configured:

- CS-RNTI (Configured Scheduling RNTI) for retransmission
- Periodicity of the configured grant Type 1
- Offset of the resource with respect to SFN (System Frame Number) = 0 in time domain
- The number of HARQ processes

Once an uplink grant is configured, the uplink grant is automatically repeated periodically so that the UE can go ahead to perform uplink data transmission without the regular scheduling hand-shake. Figure 4.13 shows the call flow of the type 1 UL GF (Grant Free) transmission.

[170] 3GPP TS 38.321 v 15.5.0, *NR Medium Access Control (MAC) Protocol Specification (Rel. 15).* March 2019.

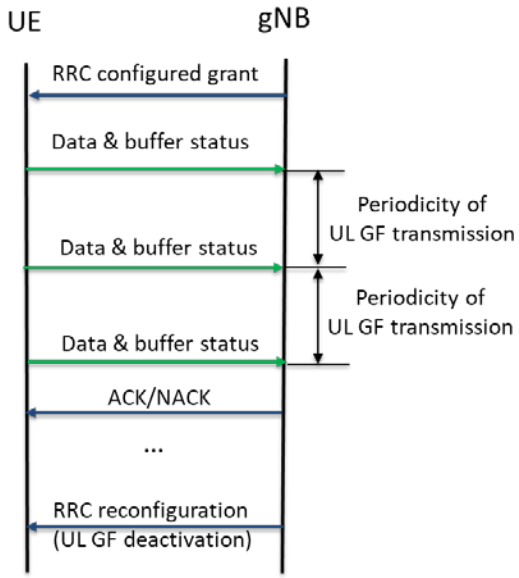


Figure 4.13. Uplink Grant Free Transmission – Type 1.

When a configured grant is released by upper layers, all the corresponding configurations will be released, and all the corresponding uplink grant will be cleared right away.

For Configured Grant Type 2, additional L1 signaling is introduced when a fast modification of semi-persistent resource allocation is needed. This enables flexibility so that UL grant free transmission can be triggered based on URLLC traffic properties such as packet arrival rate, number of UEs sharing the same resource pool and/or packet size. For Type 2, activation and deactivation of the UL GF transmission are independent among the serving cells. During Type 2 RRC configuration, the following parameters are configured:

- CS-RNTI for activation, deactivation and retransmission
- Periodicity of the configured grant Type 2
- The number of HARQ processes

Once the uplink grant free transmission is activated, the UL transmission will be triggered periodically without UL grant until the deactivation signaling is received by the UE. The configured uplink grant will be released immediately after the deactivation is triggered. Figure 4.14 shows the call flow of the procedure.

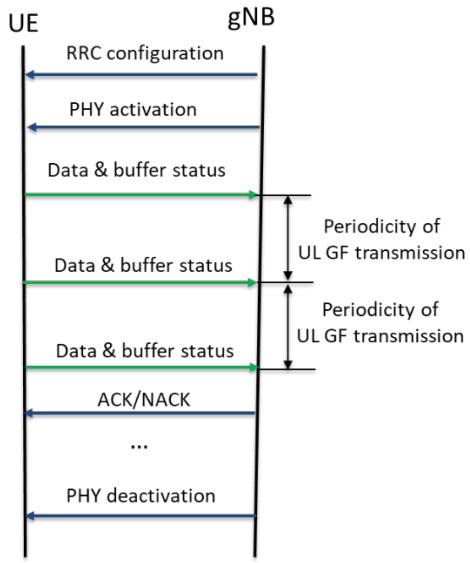


Figure 4.14. UL Grant Free Transmission-Type 2.

4.3.5 MULTIPLEXING OF URLLC AND EMBB

When both eMBB and URLLC traffic need to be scheduled, it is important to find an efficient multiplexing method so that these two types of traffic can be served with the necessary KPI requirements. Preemptive scheduling has been proposed to support dynamic resource sharing between URLLC and eMBB. [171,172,173] The benefit of this scheme is that the system allows the URLLC packets to be transmitted right away without waiting for the completion of previously scheduled eMBB transmission. Even if eMBB traffic is scheduled on all the available radio resources, URLLC data will be scheduled for transmission by taking over the part of the previously scheduled eMBB slots as soon as the URLLC data arrive at the gNB. This will greatly reduce the DL latency for the URLLC data. As shown in Figure 4.15, once the latency critical data arrives for UE#2, the on-going transmission to UE#1 is punctured immediately, and the resources at those punctured slots will be allocated for data transmission to UE#2. In this case, UE#1 will only receive the scheduled transmitted data that are partially punctured. This type of puncturing might cause degraded decoding performance of eMBB data. To compensate for the potential performance degradation, puncturing indication is introduced by 3GPP to indicate the time and/or frequency region of the impacted eMBB resources to respective eMBB UEs. To recover the decoding performance loss, an eMBB user could locate the impacted resource based on the pre-emption indication and remove the corresponding corrupted data from its soft buffer. When the Modulation and Coding Scheme (MCS) level is low (in case of low Signal-to-Noise Ratio (SNR), this procedure can effectively rescue the corrupted data of the "victim" eMBB users. However, with the increase of MCS level, the modulation order and code rate get much higher, the available error correction capability of channel codes is not enough to recover the missing data just through a preemption indication. In this case, different methods could be adopted to recover the performance loss.

One option is that the gNB could postpone the missing pre-empted part and transmit it to the victim eMBB users as a supplementary transmission after the URLLC data transmission is completed.[174] The eMBB UEs

[171] 3GPP TR 38.802, *Study on New Radio (NR) Access Technology: Physical Layer Aspects*.
[172] Zexian Li, etc, *5G URLLC: Design Challenges and System Concepts*, 15th International Symposium on Wireless Communication Systems (ISWCS), August 2018.
[173] 3GPP R1-1701663, *On DL Multiplexing of URLLC and eMBB Transmission*. February 2017.
[174] 3GPP R1-1611222, *DL URLLC multiplexing considerations.*

could use the original and postponed transmission in decoding, which will help to improve the performance of eMBB users. Alternatively, the gNB can use HARQ mechanisms to schedule the supplementary transmission, therefore, the gNB can schedule the supplementary transmission after the "victim" users perform the initial decoding based on the pre-emption information. The UEs can then combine the supplementary transmission with the previous one for decoding. This HARQ based method requires less resources compared to the first option.

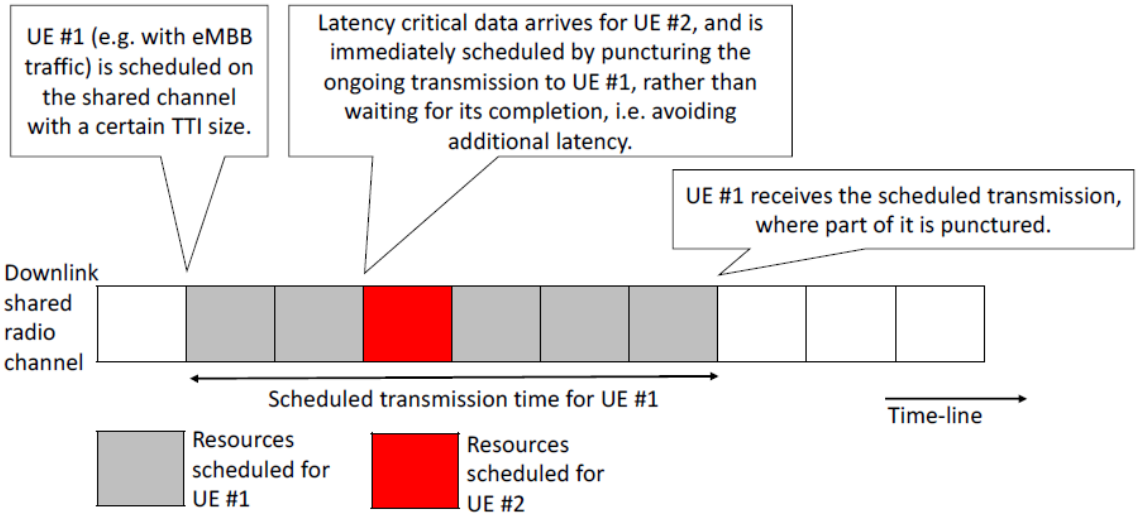


Figure 4.15. DL Multiplexing of URLLC and eMBB through Pre-emptive Scheduling.

4.3.6 ENHANCEMENTS OF PDCCH

As discussed previously, URLLC use cases have very stringent requirements on reliability. For example, some vertical applications in factory automation require a reliability of 99.9999 percent. To improve the performance of BLER (Block Error Rate) and increase reliability, 3GPP has proposed some enhancements on PDCCH, such as compact Downlink Control Indicator (DCI) and Physical Downlink Control Channel (PDCCH) repetition.[175]

In the case of compact DCI, a DCI with a smaller size will be used for URLLC applications to improve the link level performance gain of PDCCH. Specifically, the size of DCI format 0_0 (for scheduling of Physical Uplink Shared Channel (PUSCH)) and format 1_0 (for scheduling of PDSCH) will be reduced at least 10 to 16 bits compared to the DCI format in Rel-15. The size of the DCI may be reduced in the following ways:[176]

- Reduce the size of frequency domain resource assignment field in DCI formats 0_0 and 0_1 by increasing the resource allocation granularity
- Reduce the size of MCS (Modulation and Coding Scheme) field by increasing MCS index granularity. For example, use a 4-bit field in compact DCI to indicate the combined MCS index of a given MCS table
- Reduce the HARQ process number field from 4 bits to 2 bits for URLLC to support 4 HARQ processes instead of 16 HARQ processes in Rel-15. This reduction is reasonable because the low latency requirement of URLLC will require fast HARQ round trip and thus much less HARQ processes

[175] 3GPP TR 38.824 v 16.0.0, *Study on Physical Layer Enhancements for NR Ultra-Reliable and Low Latency Case (URLLC).*
[176] R1-1900803, *On potential PDCCH Enhancements for URLLC,* InterDigital Inc. January 2019.

- Reduce the size of redundancy version from 2 bits to 1 bit for URLLC since the number of allowed re-transmission is limited due to the low latency requirements of URLLC

Many simulation results show that compact DCI targeting a reduction of 16 bits compared to 40-bit Rel-15 DCI can provide 0.6 dB – 1 dB gain for AL (Aggregation Length) of 16, assuming 4GHz, 10^{-5} or 10^{-6} target BLER, 4 Rx at UE side, a CORESET (Control Resource Set) with 1 or 2 symbols in time domain and 40 MHz in frequency domain. In addition to the benefit of reliability, compact DCI also improves PDCCH resource utilization. Simulation results show that compact DCI targeting a reduction of 16 bits compared to 40-bit Rel-15 DCI can save 14 percent - 16 percent PDCCH resource used for URLLC UEs under the same assumption previously listed.

PDCCH repetition is another approach to improve the reliability of the URLLC applications. There are 5 options for the implementations:[177] [178]

- Option 1: A CORESET spans more than 3 OFDM symbols. CCE (Control Channel Element) size is extended to more than 6 (Figure 4.16)
- Option 2: A CORESET spans more than 3 OFDM symbols. CCE size is 6 but more than 16 CCEs are aggregated. (Figure 4.16)
- Option 3: PDCCH repetitions prior to PDSCH transmissions (Figure 4.17) Multiple CORSETs with the same search spaces with the same utilized CCE location are bundled
- Option 4: Independent PDCCH schedules each PDSCH repetition (Figure 4.18)
- Option 5: Multiple PDCCHs schedule PDSCH repetition with indication of the number of repetitions in each DCI (Figure 4.19)

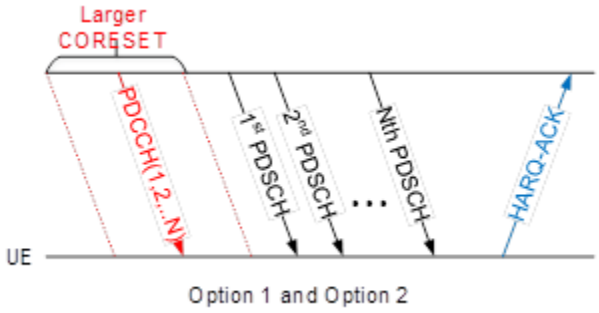


Figure 4.16. Larger CORESET for PDCCH.

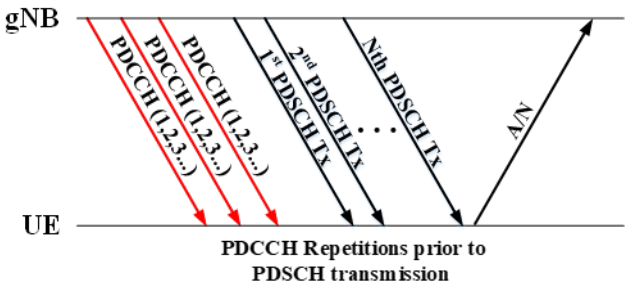


Figure 4.17. PDCCH Repetitions Prior to PDSCH Transmission.

[177] 3GPP R1-1901459, *Summary of 7.2.6.1.1 Potential Enhancements to PDCCH*. January 2019.
[178] 3GPP R1-1900399, *PDCCH Enhancements for NR URLLC.* January 2019.

Both Option 1 and 2 can provide higher reliability of PDSCH and PUSCH, but the CORESET is larger. In addition, CORESET cannot be shared with other applications. The drawbacks of these two options are the potential impacts on the existing specification. A new CORESET with larger CCE and larger CORESET length will have to be defined in the new release. Compared to Options 1 and 2, Option 3 has relatively smaller impact on the specification since there is no need to define a new CORESET. It provides better reliability of PDSCH and PUSCH and allows the CORESET for URLLC to be shared with other applications. In addition, Option 3 has precoding flexibility. The same precoding can be used in multiple CORESET to improve channel estimation, or different precoding can be used for spatial diversity. In the case of Option 4 and Option 5, there are some limitations on resource allocation flexibility. If an earlier PDCCH is missed, the UE would not know the TBS (Transport Block Size) of the initial PDSCH/PUSCH. TBS will have to be aligned for PDCCH repetitions. In other words, the same frequency allocation and the number of symbols will have to be allocated for repeated PDCCHs. To increase the resource allocation flexibility, a TBS scaling factor can be pre-configured in Option 4. Options 4 and 5 also have smaller impact on the existing specification and allow CORESET to be shared with other applications. At the time of writing, there is still no final decision in 3GPP with regards to which option will be implemented in the control channel. Further investigation is still needed to draw a final conclusion.

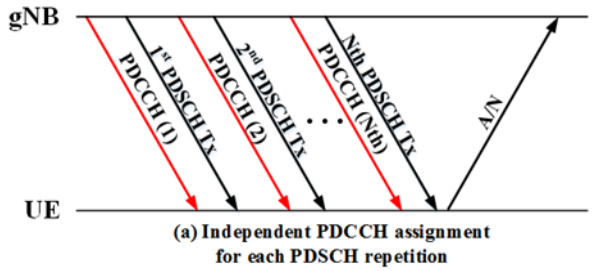


Figure 4.18. Independent PDCCH Assignment for each PDSCH Repetition.

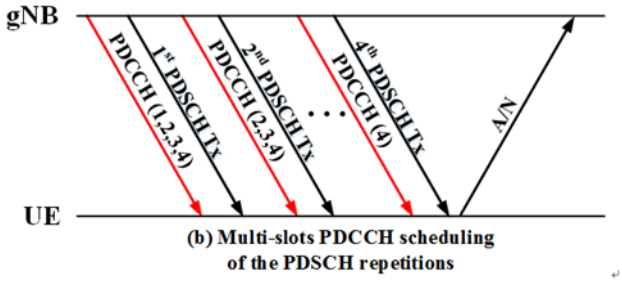


Figure 4.19. Multi-slot PDCCH Scheduling of the PDSCH Repetition.

4.3.7 ENHANCEMENTS OF HARQ FEEDBACK

In NR Rel-15, only one PUCCH is supported within a slot for HARQ-ACK transmission. To reduce latency and enable faster HARQ-ACK feedback, 3GPP proposed that more than one PUCCH for HARQ-ACK transmission within a slot should be supported in Rel-16. In the case of simultaneous eMBB and URLLC transmission, more than one PUCCH can facilitate separate HARQ-ACK feedback channels for eMBB and URLLC. In addition, 3GPP proposed in Rel-16 that at least two HARQ-ACK codebooks can be simultaneously constructed to support different service types for a UE. To enable faster HARQ-ACK feedback, Rel-16 also proposed out-of-order HARQ-ACK feedbacks. As shown in Figure 4.20, a UE can be scheduled with a URLLC PDSCH transmission after an eMBB PDSCH transmission and before the

HARQ-ACK for eMBB transmission has arrived. In Rel-15, HARQ-ACK has to be transmitted in a timely order. The UE has to transmit the first HARQ-ACK for eMBB data before it can transmit the second URLLC HARQ-ACK. The out-of-order HARQ-ACK transmission in Rel-16 allows the UE to transmit the URLLC HARQ-ACK first, even though the URLLC PDSCH data arrive after the eMBB data. This will greatly reduce the round-trip latency for URLLC.

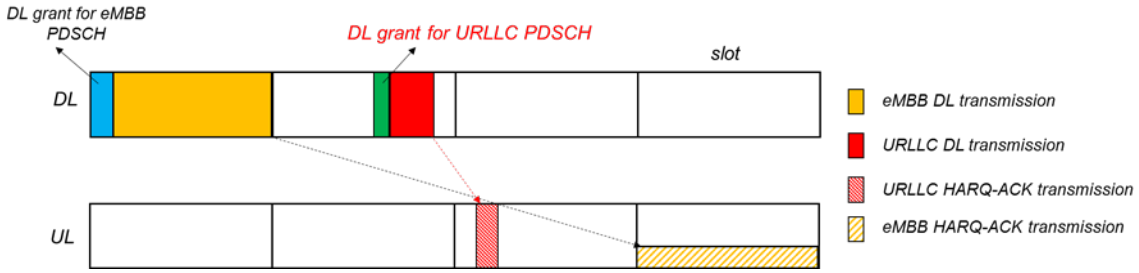
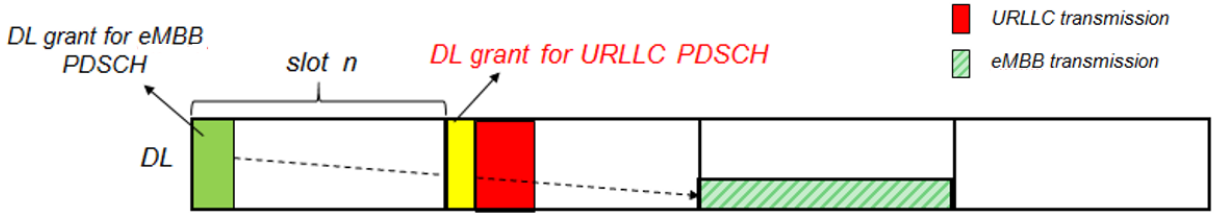


Figure 4.20. Out-of-order HARQ-ACK Feedback in Case of DL Transmission of eMBB and URLLC.

In a UE with both eMBB and URLLC services, DL grant for URLLC PDSCH transmission might come later than the DL grant for eMBB transmission. In this case, gNB should prioritize the URLLC traffic due to its stringent requirement on latency. 3GPP Rel-16 proposed out-of-order PDSCH/PUSCH scheduling to facilitate the prioritized transmission of URLLC. As shown in Figure 4.21 and Figure 4.22, URLLC transmission is scheduled before the eMBB transmission even though the DL grant for URLLC arrives later than that for eMBB.



Error! Reference source not found.**Figure 4.21. Out of order scheduling for DL URLLC transmission.**[179]

[179] 3GPP R1-1901695, *Enhancement for Scheduling/HARQ/CSI Processing Timeline.*

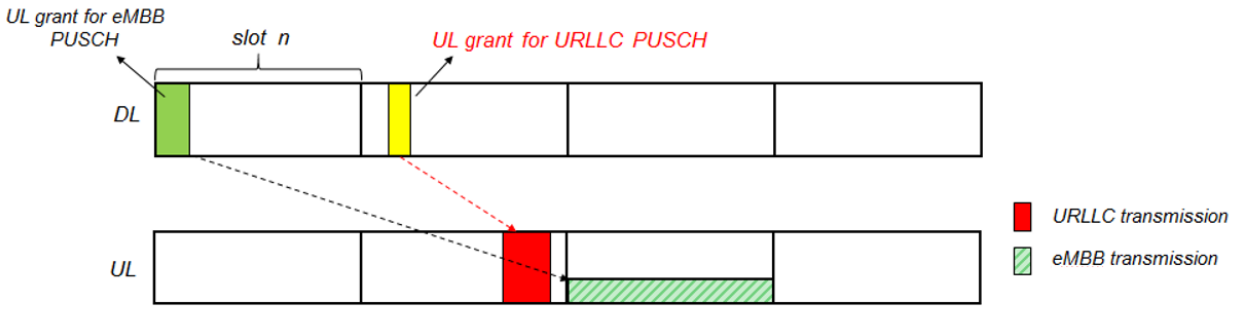


Figure 4.22. Out of Order Scheduling for UL URLLC Transmission.

4.3.8 SUPPORT OF SEPARATE CQI AND MCS TABLES FOR URLLC

To meet the stringent reliability requirements, the conservative MCS selection with lower coding rate might be used more frequently for URLLC than eMBB for 10^{-5} target BLER. The lower coding rate should be included in the URLLC MCS table. 3GPP Rel-15 defines separate MCS tables for PDSCH (Table 5.1.3.1-3 in TS 38.214) and PUSCH (Table 6.1.4.1 – 2 in TS 38.314) for this purpose. In the new tables, 5-bit MCS index is used to provide higher granularity of MCS levels. Lower coding rates together with lower spectral efficiency are included in the MCS tables. [180] [181]

[180] 3GPP 38.214 V 15.5.0, *NR Physical Layer Procedures for Data*, March 2019.
[181] 3GPP R1-1807730, *Offline Discussion on Support of Separate CQI and MCS Table for URLLC*.

Table 4.7. MCS Index Table Defined for URLLC PDSCH.

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x [1024]	Spectral efficiency
0	2	30	0.0586
1	2	40	0.0781
2	2	50	0.0977
3	2	64	0.1250
4	2	78	0.1523
5	2	99	0.1934
6	2	120	0.2344
7	2	157	0.3066
8	2	193	0.3770
9	2	251	0.4902
10	2	308	0.6016
11	2	379	0.7402
12	2	449	0.8770
13	2	526	1.0273
14	2	602	1.1758
15	4	340	1.3281
16	4	378	1.4766
17	4	434	1.6953
18	4	490	1.9141
19	4	553	2.1602
20	4	616	2.4063
21	6	438	2.5664
22	6	466	2.7305
23	6	517	3.0293
24	6	567	3.3223
25	6	616	3.6094
26	6	666	3.9023
27	6	719	4.2129
28	6	772	4.5234
29	2	reserved	
30	4	reserved	
31	6	reserved	

Table 4.8. MCS Index Table Defined for URLLC PUSCH.

MCS Index I_{MCS}	Modulation Order Q_m	Target code Rate R x 1024	Spectral efficiency
0	*q*	60/*q*	0.0586
1	*q*	80/*q*	0.0781
2	*q*	100/*q*	0.0977
3	*q*	128/*q*	0.1250
4	*q*	156/*q*	0.1523
5	*q*	198/*q*	0.1934
6	2	120	0.2344
7	2	157	0.3066
8	2	193	0.3770
9	2	251	0.4902
10	2	308	0.6016
11	2	379	0.7402
12	2	449	0.8770
13	2	526	1.0273
14	2	602	1.1758
15	2	679	1.3262
16	4	378	1.4766
17	4	434	1.6953
18	4	490	1.9141
19	4	553	2.1602
20	4	616	2.4063
21	4	658	2.5703
22	4	699	2.7305
23	4	772	3.0156
24	6	567	3.3223
25	6	616	3.6094
26	6	666	3.9023
27	6	772	4.5234
28	q	reserved	
29	2	reserved	
30	4	reserved	
31	6	reserved	

4.4 NB-IOT AND EMTC ENHANCEMENT

Although NB-IoT is most closely associated with LTE, NB-IoT enhancements involve connection to 5GC and also for coexistence with NR. These enhancements are enabling NB-IoT to become a part of 5G IoT as well.

In the case of connectionless small data DL transmission, Mobile-Terminated Early Data Transmission (MT-EDT) is supported in Rel-16. This is to improve the DL transmission efficiency as well as the UE power consumption. Use cases that require DL data transmission with or without UL data transmission as a response are supported for MT-EDT. DL small data can be transmitted over the paging message or during the paging steps, which is beneficial for the use cases where the application level acknowledgement is not expected by the device. In these cases, the application can consider the radio layer delivery information of

the paging message as successful transmission of the small data transmission to UE.[182] The quality report in MSG3 is specified at least for EDT (Early Data Transmission).

Rel-15 specified a downlink WUS (Wake-Up-Signal) being transmitted prior to a Paging Occasion (PO) to indicate whether the UE needs to read the paging information following certain WUS in a single DRX cycle. This feature allows UE to reduce detection effort of NPDCCH and paging message, therefore can reduce signaling and UE power consumption. On the other hand, other UEs which share the certain PO also wake up to read paging message unnecessarily. To solve this issue, Rel-16 proposed that the wake-up-signal (WUS) carry additional UE grouping information. UE grouping is based on at least UE ID or some function of UE ID, and the configuration of group WUS is at least signaled in SI (System Information) signaling.[183]

To improve UL transmission efficiency and UE power consumption, Rel-16 will support UL transmission in Preconfigured UL Resources (PUR) in idle and/or connected mode based on SC-FDMA waveform for UEs with a valid Timing Advance (TA). Both dedicated and shared resources can be pre-configured for UL data transmission. In case of dedicated PUR in idle mode, the UE would be provided with a PUR configuration and TA in the initial access. The pre-configured resources are configured based on UE's traffic profile, UE capabilities and CE-level, and etcetera. They can be configured either periodically or for one transmission occasion at a time. This feature allows UE to perform UL transmission without waiting for the UL grant. For both LTE-M and NB-IoT it is desirable to able to multiplex PUR UEs in the same narrowband or carrier. For per-requested PUR this is straight forward, but for legacy dynamic transmission the scheduler would have to check whether this pre-configured resource has been assigned to other UEs before granting it as a PUR resource.[184]

To better support multi-cast as well as unicast transmission for eMTC, scheduling enhancement is specified in Rel-16.[185] In the case of unicast, the following features are supported:

- Scheduling multiple DL/UL TBs (Transport Blocks) with single DCI
- The UE should only monitor one DCI size in the UE specific search space
- The possibility of scheduling multiple DL/UL TBs is configured via RRC
- The number of scheduled TBs should be dynamically indicated in the DCI
- Individual feedback for each HARQ process is supported

In the case of SC-PTM multicast, the following features are supported:

- Using one DCI to schedule multiple TBs for SC-MTCH (Single Cell - Multicast Transport Channel)
- The possibility of scheduling multiple TBs is configured and enabled per SC-MTCH via SC-PTM (Single Cell – Point to Multipoint) configuration message in SC-MCCH (Single Cell – Multicast Control Channel)
- Multiple TBs scheduling for SC-MTCH should handle backward compatibility with Rel-14 SC-PTM

[182] 3GPP R2-1900320 Further analysis on solutions for MT EDT during paging.
[183] 3GPP R1-1812454 *UE-group wake-up signal for eMTC*. November 2018.
[184] 3GPP R2-1816644 *Transmission in preconfigured uplink resources*. November 2018.
[185] 3GPP, R1-1812135 *Scheduling multiple DL/UL transport blocks for SC-PTM and unicast*. November 2018.

5. CONCLUSION

As described in Section 2 of this whitepaper, the market for IoT devices is growing and expected to continue to grow as more and more “things” are connected to the Internet. These include things used in our homes such as smart appliances and smart security, things in our communities such as smart utilities and smart parking meters, things in offices and factories such as smart printers and robotic equipment. As the number and complexity of smart things increases, the technology to support them is also being developed. We see expansion in cloud computing, edge cloud, and artificial intelligence being applied to IoT use cases. We see security enhancements to keep all the data generated by these “things” secure from hackers and other security breaches. V2X, drone control and robotic controls pose different more stringent requirements and we see a need to support mobile as well as stationary IoT devices. 3GPP technology provides the flexibility to support the many diverse requirements of the IoT.

From the marketing analysis, it’s easy to see the growth path for IoT – in our homes, in our communities, in the office and industrial settings. More bandwidth, more efficient use of resources, and wider areas of coverage are needed to meet the projected growth of IoT over the coming years. A flexible system, such as 5G, can provide the infrastructural support for the expanding demands coming in the form of not only more devices but also specialized devices with varying performance requirements.

3GPP is meeting the challenge of supporting the IoT with many enhancements in 5G. Building on the solid base of IoT support provided in 4G with eMTC and NB-IoT, 5G focuses on supporting the ever-increasing number of IoT devices as well as the specific KPIs (for example, low latency, high reliability, positioning accuracy) necessary for Industrial, Enterprise, and even Consumer use of IoT devices. The need for URLLC and TSN functionality is addressed in both the core network and radio interface. Similarly, many resource efficiencies are implemented to support massive numbers of IoT devices with varying requirements for data transmission, mobility, and accessibility. Standard approaches to supporting non-public networks allow use of 5G in industry and enterprise environments, while network slicing allows customizability of network resource deployments to meet specific use case requirements. Enhancements to spectrum options for licensed, unlicensed, and shared bands expand the use of 5G technology to new arenas.

APPENDIX

APPENDIX A. ACRONYMS

1G First Generation

2G Second Generation

3G Third Generation

3GPP Third Generation Partnership Project (Global Standards Organization)

4G Fourth Generation

5G Fifth Generation

5G-ACIA 5G Alliance for Connected Industries & Automation

5GC 5G Core

5G-LAN-VN 5G-Local Area Network-Virtual Network

ACK Acknowledgement

AESE Architectural Enhancements for Service Capability Exposure

AI/ML Artificial Intelligence/Machine Learning

AL Aggregated Length

API Application Programming Interface

AR Augmented Reality

ARC Automatic Retransmission Request

A/V Audio/Visual

BL Bandwidth-reduced Low-complexity

BLER Block Error Rate

CA Carrier Aggregation

CAGR Compound Annual Growth Rate

Cat Category

CE Coverage Enhancement

C-IoT Cellular IoT

CORESET Control Resource Set

CN Core Network

CPS Cyber-Physical System

CS-RNTI Configuration Scheduling-Radio Network Temporary Identifier

CUPS Control and User Plane Separation

dBm Decibel/milliwatt

DCI Downlink Control Indicators

DDoS Distributed Denial of Service

DÉCOR Dedicated Core Network

DL Downlink

DMRS Demodulation Reference Signal

DPDK Data Plane Development Kit

DN Data Network

DNN Data Network Name

DRX Discontinuous Reception

DVR Digital Video Recorder

EAB Extended Access Bearer

EB Exebytes

EC-GSM-IoT Extended Coverage- GSM-IoT

EDT Early Data Transmission

eDRX extended Discontinuous Reception

eHPLMN Equivalent Home Public Land Mobile Network

eMBB enhanced Mobile Broadband

eMTC Enhanced Machine Type Communication

EPC LTE Core

ERP Enterprise Resource Planning

gNB Next Generation NodeB

FDD Frequency Division Duplex

FD.io Fast Data Input/Output Project

GB Gigabit

GF Grant Free

GHz Gigahertz

GNSS Global Navigation Satellite System

GPRS General Packet Radio System

gPTP generic Precision Time Protocol

GROUPE Group based Enhancements

GSM Global System for Mobility

GTP-U GPRS Tunneling Protocol User Plane

HARQ Hybrid Automatic Retransmission (or Repeat) Request

HLCom High Latency Communications

HRLLC Highly-Reliable Low Latency Communications

HSPA High Speed Packet Access

HPLMN Home Public Land Mobile Network

IAB Interactive Advertising Bureau

IEEE Institute of Electrical and Electronics Engineers

IIC Industrial Internet Consortium

IIoT Industrial IoT

IMS IP Multimedia Subsystem

IMSI International Mobile Subscriber Identity

IMT International Mobile Telecommunications

IMU Inertial Measurement Unit

IOT Internet of Things

IP Internet Protocol

IPv6 IP version 6

I-UPF Intermediate User Plane Function

Kbps Kilobits per second

km Kilometer

KPI	Key Performance Indicator
kHz	Kilohertz
LAA	License-Assisted Access
LOB	Line-of-Business
LOS	Line-of-Sight
LPWA	Low Power Wide Area
LTE	Long Term Evolution
LTE-M	Long Term Evolution for Machines
LTE-U	LTE in Unlicensed (spectrum)
KPI	Key Performance Indicator
M2M	Machine-to-Machine
MAC	Medium-Access Control
M-CORD	Mobile Central Office Re-architected as a Data Center
METIS II	Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society – II
MIoT	Massive Internet of Things
MMTel	Multimedia Telephony Service
mmWave	millimeter Wave
MPS	Multimedia Priority Service
MTC	Machine-Type Communication
MT-EDT	Mobile-Terminated Early Data Transmission
NB-IOT	Narrow Band IoT
NF	Network Functions
NGC	Next Generation Core
NG-IC	National Ground Intelligence Center
NG-RAN	Next Generation Radio Access Network
NHTSA	National Highway Traffic Safety Administration
NOMA	Non-Orthogonal Multiple Access

NR	New Radio
NPDCCH	Narrowband Physical Downlink Control Channel
OFDM	Orthogonal Frequency Division Multiplexing
OPC UA	Object Linking and Embedding for Process Control Unified Architecture
ORAN	Open Radio Access Network
OS	Operating Systems
OVS	Open V-Switch or Open Virtualized multi-layer Switch
P2P	Peer-to-Peer
PDCCH	Physical Downlink Control Channel
PDU	Protocol Data Unit
PHY	Physical layer
PLC	Programmable Logical Controller
PLMN	Public Land Mobile Network
PO	Paging Occasion
Prose	Proximity Services
PRACH	Physical Random-Access CHannel
PSA	PDU Session Anchor
PSM	Power Saving Mode
PUR	Preconfigured UL Resources
PUSCH	Physical Uplink Shared CHannel
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RRC	Radio Resource Control
RSN	Redundancy Sequence Number
RSRP	Reference Signal Received Power
RSU	Roadside Unit

RTT	Real-Time Text
Rx	Radio transmission
s	Second
SBA	Services-Based Architecture
SC-FDMA	Single Cell – Frequency Division Multiple Access
SC-MCCH	Single Cell – Multicast Control Channel
SC-MTCH	Single Cell - Multicast Transport Channel
SC-PTM	Since Cell – Point to Multipoint
SCS	Sub-Carrier Spacing
SI	System Information
SLA	Service Level Agreement
SMF	Session Management Function
SMS	Short Message Service
S-NSSAI	Single Network Slice Selection Assistance Information
SPS	Semi-Persistent Scheduling
TA	Timing Advance
TB	Transport Blocks
TBS	Transport Block Size
TDD	Time Division Duplexing
TIP	Telecom Infrastructure Project
TSN	Time Sensitive Networking
TTI	Transmit Time Travel
UAV	Unmanned Aerial Vehicle
UE	User Equipment
UE ID	User Equipment Identifier
UL	Uplink
UPF	User Plane Function
URLLC	Ultra-Reliable Low Latency Communication

µs	Microsiemens
USDOT	U.S. Department of Transportation
USIM	Universal Subscriber Identity Module
V2I	Vehicle-to-Infrastructure
V2X	Vehicle-to-Everything
V2V	Vehicle-to-Vehicle
VLAN	Virtual Local Area Network
VOD	Video-on-Demand
VoLTE	Voice-over-LTE
VPN	Virtual Public Network
VPP	Vector Packet Processing
VR	Virtual Reality
WAN	Wide Area Network
WLAN	Wireless Local Area Network
WUS	Wake-Up-Signal
XR	Extended Reality
ZVEI	German Electrical & Electronic Manufacturing Association

APPENDIX B. LIST OF REFERENCED 5G AMERICAS WHITEPAPERS

1. *The Status of Open Source for 5G*, 5G Americas white paper, February 2019.
2. *5G Spectrum Vision*, whitepaper by 5G Americas. February 2019.
3. *LTE to 5G – The Global Impact of Wireless Innovation,* Rysavy Research, 5G Americas, October 2018.
4. *Wireless Technology Evolution: Transition from 4G to 5G*, *3GPP Releases 14 to 16,* 5G Americas Whitepaper. October 2018.
5. *The Evolution of Security in 5G,* 5G Americas Whitepaper. October 2018.
6. *Cellular V2X Communications Towards 5G*, white paper by 5G Americas. March 2018.
7. *Wireless Technology Evolution Towards 5G*, 5G Americas Whitepaper. February 2017.
8. *LTE Progress Leading to the 5G Massive Internet of Things*, 5G Americas Whitepaper, December 2017.
9. *LTE and 5G Technologies Enabling the Internet of Things,* 5G Americas white paper. December 2016.

ACKNOWLEDGEMENTS

The mission of 5G Americas is to advocate for and facilitate the advancement of 5G and the transformation of LTE networks throughout the Americas region. 5G Americas is invested in developing a connected wireless community for the many economic and social benefits this will bring to all those living in the region.

5G Americas' Board of Governors members include AT&T, Cable & Wireless, Ciena, Cisco, CommScope, Ericsson, Intel, Kathrein, Mavenir, Nokia, Qualcomm Incorporated, Samsung, Shaw Communications Inc., Sprint, T-Mobile USA, Inc., Telefónica and WOM.

5G Americas would like to recognize the significant project leadership and important contributions of project co-leaders Betsy Covell from Nokia and Rajat Prakash of Qualcomm, as well as Vicki Livingston of 5G Americas along with many representatives from member companies on 5G Americas' Board of Governors who participated in the development of this white paper.