The Future of 5G on the Factory Floor

Rob Lodesky
Director, Strategic Accounts
HMS Industrial Networks

Presented at the ODVA
2023 Industry Conference & 22nd Annual Meeting
October 18, 2023
El Vendrell, Spain

Abstract

5G cellular connectivity brings exciting promise to the Industrial Automation sector. This paper tries to answer the questions of what enhancements to industrial communication we can expect in the near term, such as release 17 and 18, and what we can expect in the long term, such as 6G. This paper also tries to understand what near term and long term is, when we can expect technology to adopt these solutions, as well as define some of the overarching political and market fundamentals which change when and why this technology is used. We will also look at the present state of 5G solutions to drive an understanding of what forces are currently resisting this change and who is currently benefiting. This paper is a supplement to David Brandt's and Scott Griffiths paper from the 2020 ODVA conference "5G - Not Just for Cell Phones Anymore."

The certain path for 5G on the industrial floor, paved by 5G-ACIA, the supplementary organization to assist 3GPP in delivering 5G technology for industrial applications, has had a long gestation period since ACIA's inception. What can we expect in the next generation of 5G releases, and upcoming 6G technology on industrial networks?

Keywords

URLLC
ACIA
UE
SCS
IP
gnB

Release 16

There have been significant changes in release 16 to the main tenets of industrial space, specifically enhancing, or, increasing the flexibility of the URLLC specification to allow for further use cases to be applied. With release 15, a base foundation was added so as the reliability (99.9999% or 6 9's) and latency (1-2 ms) requirements of URLLC could be achieved within release 16. The foundational elements were many. Use of multiple SCS and Bandwidth flexibility are used to create dynamically arranged Grants, which allow for per traffic scheduling. This allows for per traffic scheduling, instead of the cell based per unit (per UE) scheduling. In addition, per traffic schedule also segments types of traffic, i.e., status updates over parameter updates from a downlink and uplink perspective. Within this solution resides the problem of allocating resources for time sensitive data vs non-time sensitive date which share the same resource. This problem has been addressed in release 16 by creating a priority of priorities, as it were by addressing priority by normalized use of the resource and available data to transmit. Reliability constraints with release 15 regarding lack of a proper compression solution to lower radio transmission have been addressed with a new compression algorithm, which does not compress the IP or transport layers, thus allowing Industrial networks to properly utilize the ethernet frames. To address resiliency, packet duplication would allow multiple gnB (antennas) to address a node at once.

Scheduling Enhancements

Unlike 4G LTE solutions, 5G allows for much higher flexibility and scalability in scheduling uplink and downlink methods to decrease round trip time, or RTT (9). Instead of connecting directly to a system with one channel, a System can connect to multiple points within the system with their own dedicated schedule, thus optimizing the complete network.

Increasing the PDCCH (Downlink transmission) Monitoring bandwidth

In release 15, downlink transmission channels can only be monitored by slot. This prohibits accurate monitoring due to sub-slot uses using OFDM. Monitoring errors can occur, which can increase latency and reliability possibilities. Release 16 monitors downlink transmission in a different way. The downlink transmission is monitored per slot symbol, and the number of downlink transmission frequencies are limited which means transmission data is constantly applied.

Continued Challenges

Thus far, we have looked at the current state of 5G, now we delve into the challenges running a Industrial network over 5G, and the timeline for significant enhancements. One topic of concern is ensuring an IP routable network, such as that which 5G is, can be used. There are multiple protocols, such as DLR, which are based on Layer 2 implementations, which do not have IP frames, thus, are not routable in a 5G network. In addition, non-LLDP devices use a discovery protocol using UDP/IP, the protocol uses IP Broadcasting, which is not IP routable. In addition, IP Multicast frames are routable on an IP network, but need routers with specialized configuration.

Most of these functions are not necessary on every application, therefore, they can be ignored or turned off on the Controller or device. This may cause problems due the inherent default options for most devices are DLR and discovery.

One option which would allow both DLR and discovery to occur is use of Ethernet Packet Data units (PDU). Although PDU's are a relatively old technology, its use on the Ethernet layer (Layer 2) is a new interpretation of this technology and is not of general use in the market yet. Additionally, significant use of the broadcast protocols may cause saturation of the cellular connection. Additionally, LAN switches typically implement IGMP snooping to handle excessive packet congestion by limiting multicast frames to the subscribed IP subscribers. IGMP is an IETF (Internet Engineering Task Force) recognized protocol. IEEE (Institute of Electric and Electronic Engineers) does not offer IGMP as a feature set. IEEE is recognized by the organization which codifies and standardizes current cellular network technology, 3GPP (3rd Generation Partnership Project) as a partner in standardization body. As such, IGMP may not be considered as a protocol adopted by the telecom industry, therefore, may not be supported by cellular networks.

One workaround to consider is use of tunneling on devices on the 5G network implementing EtherNet/IP. This would allow use of all Layer 2 protocols, as well as broadcast messages. This can be done by Virtual Extensible Local Area Network implementation (13) or Generic Routing Encapsulation (GRE). This provides scalability and significant segmentation implementations heretofore unavailable. This implementation does have some challenges. A 64-bit header is necessary for VXLAN, therefore more overlay and processing therein, however advanced QoS implementation would lower the degree to which this degrades the determinism. VXLAN behavior on a 5G network has been studied (14). In addition, VXLAN is a IETF standard protocol, IEEE has not investigated standardizing VXLAN technology. Generally, 3GPP has looked at IEEE as the de facto organization to align internet technologies.

TSN Enablement

Enablement of 1588, as well as TSN is an important function which enhances utilization of 5G on the factory floor. As outlined in the 5G-ACIA white paper “Integration of 5G with Time Sensitive Network for Industrial Communications (10). Release 16 adds (g) PTP synchronization which allows IEEE 1588 operation. Motion applications seem ideal for 5G use, as many of the incongruent variables, such as high node count and unknown interference do not exist in many highly deterministic applications, such as CNC’s and Semiconductor fabrication processes.

Release 17

RedCap Enhancements

RedCap, or reduced capacity, is built for solutions which do not need the determinism of URLLC requirements. Generally, increased data throughput without a significant increase in cost. This opens the possibility for 5G wearable devices and powered industrial sensors for monitoring vibration, pressure, and temperature. These types of sensors do not require high latency but high reliability, which RedCap can efficiently deliver.

The RedCap spec was devised to ensure that heterogeneous 5G solutions were available. One of the significant advantages of RedCap is designing 5G to a more power-efficient solution, to allow battery power. Fast upload and download speeds create a significant burden on the power. Thus, some creativity can be derived to lessen the load on the processor or change the processor altogether to a lower-style version.

Additionally, RedCap allows for higher latency for non-deterministic applications, making it more flexible and versatile than traditional 5G solutions. Of course, latency reduction allowance is the byproduct of various changes in specification. Halving the bandwidth, allowing half-duplex, limiting the number of Quadrature Amplitude modulation channels, and limiting the number of transmit and receive channels create this relaxed latency.

Another critical advantage of RedCap is its cost-effectiveness. eMBB requirements place a high cost on video upload and download times, which require an increase in gating, and clock speed, among other processor enhancements. Wearables are a significant discount; just the chips cost will be inherently 5 to 6 times less expensive to meet the current market pricing (12). This makes it an attractive option for companies looking to develop wearable devices and sensors that are more affordable for consumers.

mmWave Enhancements

Release 17 enhances the mmWave spectrum, 24 GHz to 300 GHz range, called FR2-2, to use the 52.6 to 71 GHz range. The mmWave spectrum is in a higher part of the spectrum, which limits its use to closer range. It does so by use of beamforming technology, the range capacity is significantly increased compared to your standard 3.5 GHz range. Beamforming can be used in several ways to improve coverage, network efficiency, and reduce interference, making 5G networks even more reliable than before.

The FR2-2 update adds almost 18 GHz in new bandwidth, which is a significant increase. And since the bandwidth is within a range of usable high frequency, it's perfect for places like stadiums, docks, train depots, and large factory floors. The FR2-2 5G NR specification provides enhanced features for use cases that require ultra-low latency and ultra-high throughput. This means that it includes features such as frequency division duplex (FDD) and time division duplex (TDD) operation, support for Massive MIMO (Multiple Input Multiple Output) antenna arrays, beamforming, and beam tracking. It also supports multiple access technologies, such as Orthogonal Frequency Division Multiple Access (OFDMA) and Non-Orthogonal Multiple Access (NOMA).

One of the most exciting use cases for the higher mmWave spectrum technology is Autonomously Guided Vehicles (AGVs). AGVs require a lot of effort to engineer, commission, and maintain due to the amount of control done in the vehicle itself. By creating a pathway to low latency data, a lot of the control can be done externally, making the AGV more agile and cost-effective. The distance AGVs typically go is within the range of a mmWave antenna and switching can be coordinated to allow for switching over long distances.

Release 18

We can consider Release 18 a study release, meaning many Factory automation requirements, validations, testing, or case studies are considered and studied, rather than applied. There are significant studies under consideration. Increases in positioning for use in Factory Automation is considered practically in the Technical Specification Group Services and System Aspects under “Service requirements for cyber-physical control applications in vertical domains”. This requirements document is significant to further specification alignment to the factory automation and process industries.

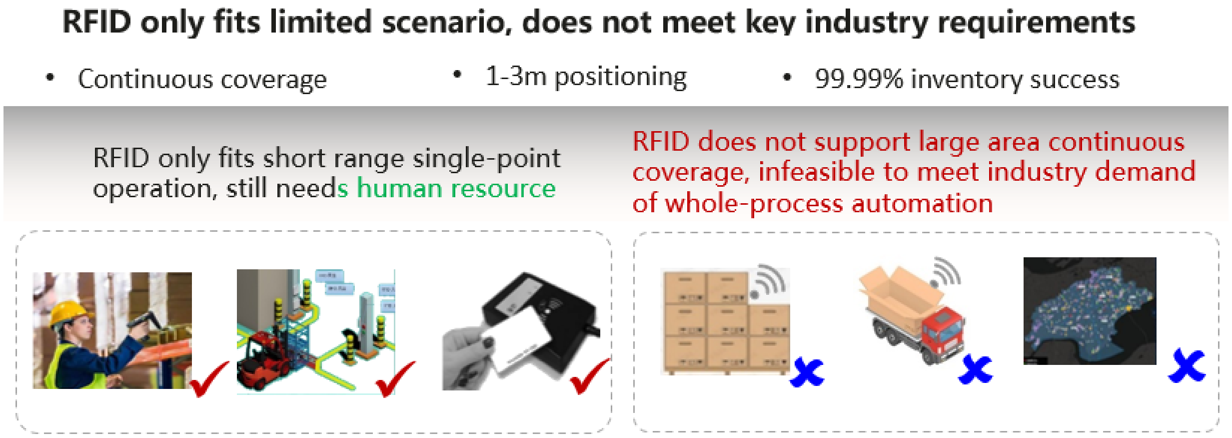
Ambient IoT

The Next generation of Low/No Power sensors and solutions will use power from various sources, known as Energy harvesting. Various considerations for applications are considered. Intralogistics and automotive inventory tracking, as well as sensing, positioning and IO could all be considered in this technology. Low cost, low/no power solutions would enable large arrays of sensors. Enabling systems which typically used RFID could enable a significant increase in data transmission.

3 categories of devices are considered.
Device A is the No Power passive device which uses backscattering transmission to react to inputs.
Device B is the Semi-passive device with some energy storage.
Device C is an active device with independent output signaling.

4 Topologies are considered when applying a theoretical Ambient IoT application within a 5G network.
Topology 1: Backscatter network to the Ambient IoT device (Device A)
Topology 2: backscatter network to an intermediate (mesh) node to the Ambient IoT Device (Device A)
Topology 3: Backscatter network to a node assistance to an Ambient IoT Device (Device B) transmitting to a Backscatter network.
Topology 4: User Equipment (5G enabled phone) to an Ambient IoT Device (Device A)

Urgency is needed due to multiple considerations from other technologies, such as Wi-Fi, Bluetooth, and many others. Some estimations of possible applications are like the RFID market. RFI currently at 20 billion units in 2022, should reach 49 billion Units by 2031 (14). Considerations to limit topologies and devices into release 19 would increase the speed of the standardization. This was done with 5G specifications with emBB releases early in the scheduling, with URLLC happening much later. To achieve some result, case studies in TR 22.840 and outlined basic use cases, as inventory is, and a focus towards these baseline attributes, such as communication range, positioning accuracy, etc.



Source: Global RFID Sensor Market, Global Industry Analysis, Size, Share, Growth, Trends and Forecast, 2021–2031

One technology to consider this approach would be cross-band Operation. Cross-band operation eliminates one of the biggest power consumers on a chip, the crystal oscillator from the BOM and, instead, transmits a clock reference signal. In industrial communication, these sensors would theoretically be synchronous to rest of the network. More studies need to be completed to understand the baud rate to which this synchronicity is possible. As well, what bands should be used is currently being discussed, with some focus on the frequency division duplex, or the 700-2000 MHz, known as FDD spectrum. The low frequency bands are considered due to the nature of low power.

Release 21 (6G)

The promise of 6G is currently being formulated by 3GPP, research starting in 2024 under release 21, with specs starting in 2028 and roll out expected mid-2029. Transceiver Demonstrators, such as those capable of Sub THz, are being developed by various institutes around the world, in anticipation of 6G ramp up.

One of the latent pains not considered in release 19 or 20 is the ability to apply deterministic systems reliably with 5G. and 6G is determinism End to End, as outlined in paper: "Towards Deterministic Communications in 6G Networks" 1 ms is not and will not be a deterministic cycle time for an end-to-end system. Truly, 5G networks and the URLLC requirement only solves some, although not all, applications in the factory automation spectrum.

The argument is harnessed by addressing both TSN (an IEC 802) and IETF (Internet Engineering Task Force) Deterministic Networking protocols into what 5G is capable of.

The fundamental weakest link is reliability compared to wired construction, as the Radio Access Network (RAN) is unquantifiable in its complexity. Significant work needs to be completed with Frame Replication and elimination for reliability, an 802.1CB specification, to consider the enhancements this creates. As well DetNet PREOF (Packet Replication, Elimination and Ordering function) could enhance the reliability as well.

Another problem is Admission control (how do you ensure new users do not impact the determinism of the network) and resource management (How do you ensure efficient use of transmission capabilities) on a 5G network.

In the end, 5G cannot match wired systems regarding bandwidth and low error rates (it can match one, but not the other), regardless of test bed examples. However, there are ways around this implicit problem. As discussed previously, DetNet and TSN have ways of diversifying the network paths to enhance reliability.

Random influences, such as interference based on mobility or non-static real life network models (channel effects?), have held experimental solutions from production. However, there are theoretical solutions based on information or queue enhancements which seem to create guarantees data and service arrival patterns.

For example, a study by IEEE shows "Wi-Fi 6 can support <; 1 ms applications at a very low load, whereas the performance gap with respect to 5G NR reduces as delay requirements are relaxed to 10-100 ms."

THz spectrum use

One of the fundamental changes in release 21 (6G) and all previous generations of Telecom communication is the use of the Terahertz spectrum, which starts at 100 GHz. 5G, for example, uses ranges of the Gigahertz spectrum, limited to 71 Gigahertz. Some major differences in using this spectrum are its theoretical capacity to move more data per second. However, this comes with some downside, as the wavelength noise constraints limit its range. 5G saw this with FR2, which used the 24 to 52 GHz range. Some of these frequencies currently focused on are the waveguide D band 110−170 GHz, due to its past use in television broadcasting (1) as well as IEEE defined 252−322 GHz. Verizon. Alternatively, the FCC has also allocated non-THz ranges to be used with 6G, such as the 7-16 GHz spectrum.

Antennas which can pick up these signals are at its infancy, with studies showing significant hardship in the transmission of these signals using standard methods, although new embedded design focus should manage to deliver usable methods on bother the receive and transmit side (3,4).

xURLLC

Organizations, such as ACIA (5) , as well as Universities (6), have commented on URLLC having a minimum latency of 1 ms, which, in many factory applications, is sub-optimal. In addition, many KPIs for Industrial Automation, such as high spectrum efficiency, throughput, energy efficiency, network availability, jitter, & round-trip delay (7) remain bottlenecks which prohibit large scale use in industrial automation. Next Generation URLLC, or xURLLC, aims to anticipate addressing these KPI’s to deliver next generation applications, such as sub millisecond SCADA systems, Cloud controllers. xURLLC is currently in study phase in 3GPP, COMSec and many other organizations.

References (optional)

(1) https://www.nature.com/articles/s41467-023-36621-x
(2) https://circleid.com/posts/20230531-fcc-touts-6g
(3) https://ieeexplore.ieee.org/abstract/document/9124764
(4) https://ieeexplore.ieee.org/document/9147140
(5) https://5g-acia.org/wp-content/uploads/5G-ACIA_WP_Key-5G-Use-Cases-and-Requirements_SinglePages.pdf
(6) https://arxiv.org/abs/2304.01299
(7) https://www.comsoc.org/publications/journals/ieee-jsac/cfp/xurllc-6g-next-generation-ultra-reliable-and-low-latency
(8) https://5g-acia.org/whitepapers/our-view-on-the-evolution-of-5g-towards-6g/
(9) https://ieeexplore.ieee.org/cart/download.jsp?partnum=9356508&searchProductType=IEEE%20Journals%20Magazines
(10) https://5g-acia.org/whitepapers/integration-of-5g-with-time-sensitive-networking-for-industrial-communications/
(11) https://www.3gpp.org/technologies/ind-5g
(12) Specification #: 23.501: portal.3gpp.org
(13) https://datatracker.ietf.org/doc/html/rfc7348
(14) Satyanarayana, Mahesh, "VXLAN EXTENSIONS FOR 5G USER EQUIPMENT SESSIONS", Technical Disclosure Commons, (May 13, 2022) https://www.tdcommons.org/dpubs_series/5134

(15) “Towards Deterministic Communications in 5G Networks: State of the Art, Open Challenges and the way forward” Gourav Prateek Sharma, et al
****