How Should the 5G-NR Standard Evolve to Address Massive IoT Applications?

SEQUANS

Contents

1. Introduction

Ever since the dawn of 2G, cellular has proven to be an ideal connectivity solution for IoT applications. Reasons for this are many: 1) it is reliable (maintained by mobile network operators), 2) it is secure (based on standards), 3) it is available in all regions worldwide, and 4) it is scalable.

If implemented correctly, and based on today's 4G and 5G standards, cellular IoT enables cost and power-optimized solutions for all types of IoT devices. Low cost and low power are prerequisites for IoT to realize sustainability and reach massive adoption.

3GPP has done the necessary work to define the various device categories dedicated to supporting IoT. These are: LTE Cat 1 (Release 8), LTE-M and NB-IoT (Release 13), and a cost-optimized variant of LTE Cat 1, called Cat 1bis (Release 14). Although these device categories cannot be defined solely by their targeted data rates, there is an approximate mapping between these LTE categories and the speed at which the application data can be transferred in and out of the devices. For NB-IoT this speed is a few tens of kbps, for LTE-M it's a few 100s of kbps, and for Cat 1 or Cat 1bis, it's a few Mbps.

Although there is no clear split on the mapping of LTE technologies to IoT applications, we can still see some trends of applicability as illustrated below. Higher capacity devices are addressed with more capable LTE categories, allowing speeds of a few 10s of Mbps up to a few 100s of Mbps with LTE Cat 4 and above. See Figure 1.

Cellular modules dedicated to LTE Cat 1 are priced, depending on the number of carriers supported, in the high-teens range; LTE Cat 1 bis modules are priced in the low teens; and LTE-M and NB-IoT modules are expected to be in the single digits. See Figure 2.

Figure 1: IoT Applications by Category

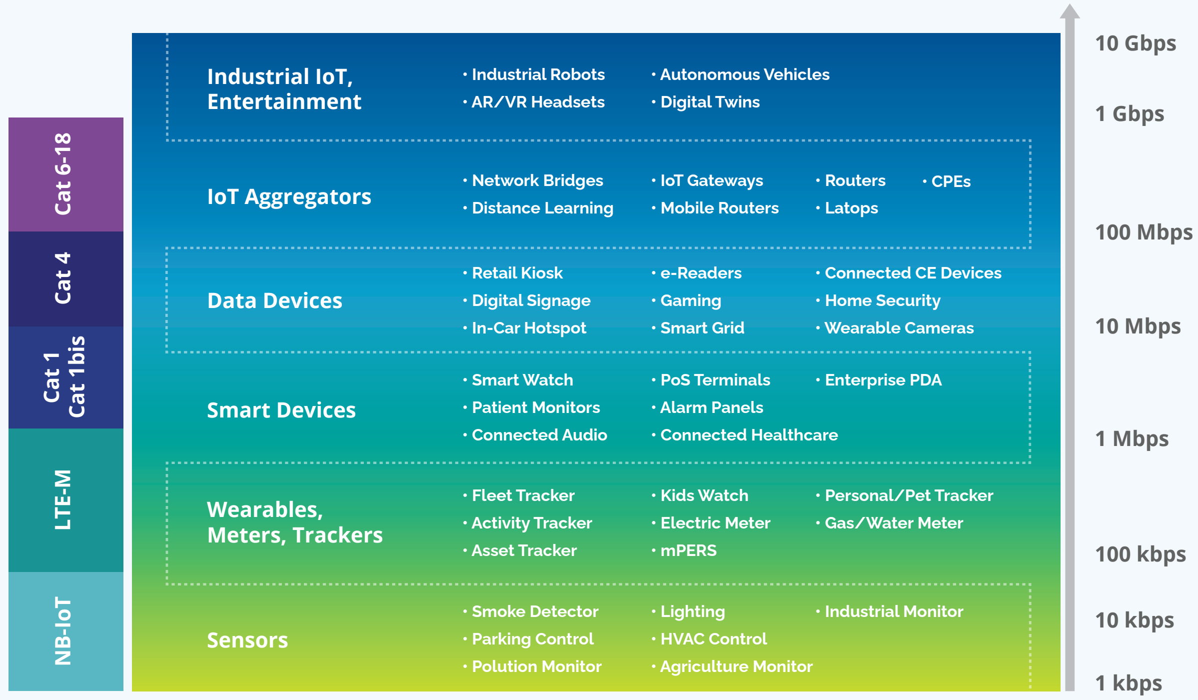
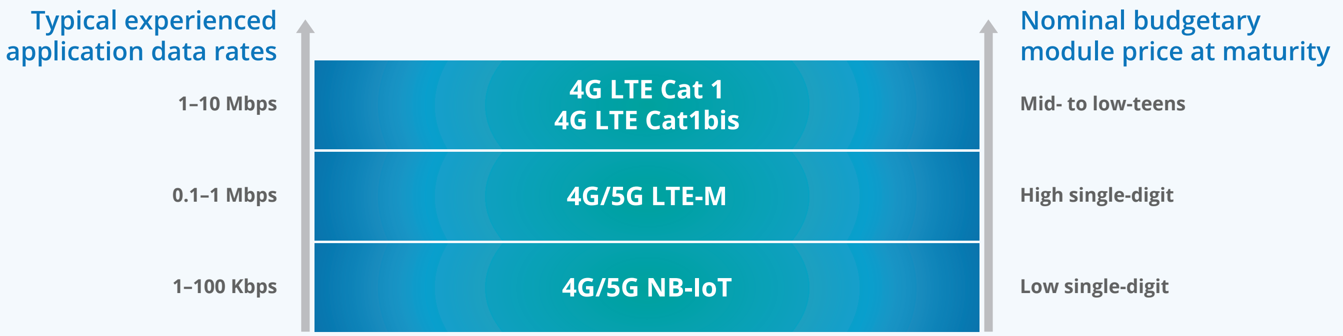


Figure 2: Cellular IoT Module Pricing



In terms of power consumption, an optimized solution can go all the way down to 1µA in a dormant state (called PSM, or Power Saving Mode), to a few 10s of µA in a reachable state (called eDRX, or extended Discontinuous Reception), to a few mA in an idle state (called DRX, or Discontinuous Reception), and a few 100mA in an active state (receiving or transmitting data).

As we see, with 4G LTE, we need at least 2 LTE categories—and dedicated implementations—to address massive IoT applications with the right tradeoff in terms of performance, price, and power consumption, namely LTE Cat 1bis or Cat 1, and LTE-M or NB-IoT.

Consequently, when we move to 5G-NR (New Radio), we should plan for similar variants of 5G-NR low device categories if we need to address all massive IoT applications with 5G-NR and prepare for a smooth transition of cellular networks from 4G to 5G.

2. Benefits of 5G-NR for IoT

When 5G was first defined, it aimed at addressing the increasing capacity needs of the networks and the exponentially increasing data usage of smartphones. It relies on new technologies that fall under the umbrella of New Radio, or 5G NR.

The first objective of 5G in implementation was to address the enhanced Mobile Broadband (eMBB) requirements that would enable a few Gbps of application data rate. And eMBB is one angle of the triangle (see Figure 3) that defined the scope of 5G. This was the priority in 3GPP Release 15.

There are other angles to 5G that address applications including ultra-reliable Low Latency Communications (urLLC) for medical, gaming, industrial, and banking; and massive Machine Type Communications (mMTC). While eMBB and urLLC were addressed initially with 5G NR (starting with Release 15 and evolving to Release 16), it was decided that mMTC would be addressed with 4G LTE-M and NB-IoT.

With this approach, the objectives of 5G were achieved, resulting in a very efficient and scalable technology with core network evolutions that enabled very high capacity and high individual data rates with optimized latencies for critical applications; and including re-use of the legacy LTE-M and NB-IoT technologies for the cost-optimized, low-power needs of the massive market of connected sensors, meters, security panels, telematics, and wearables.

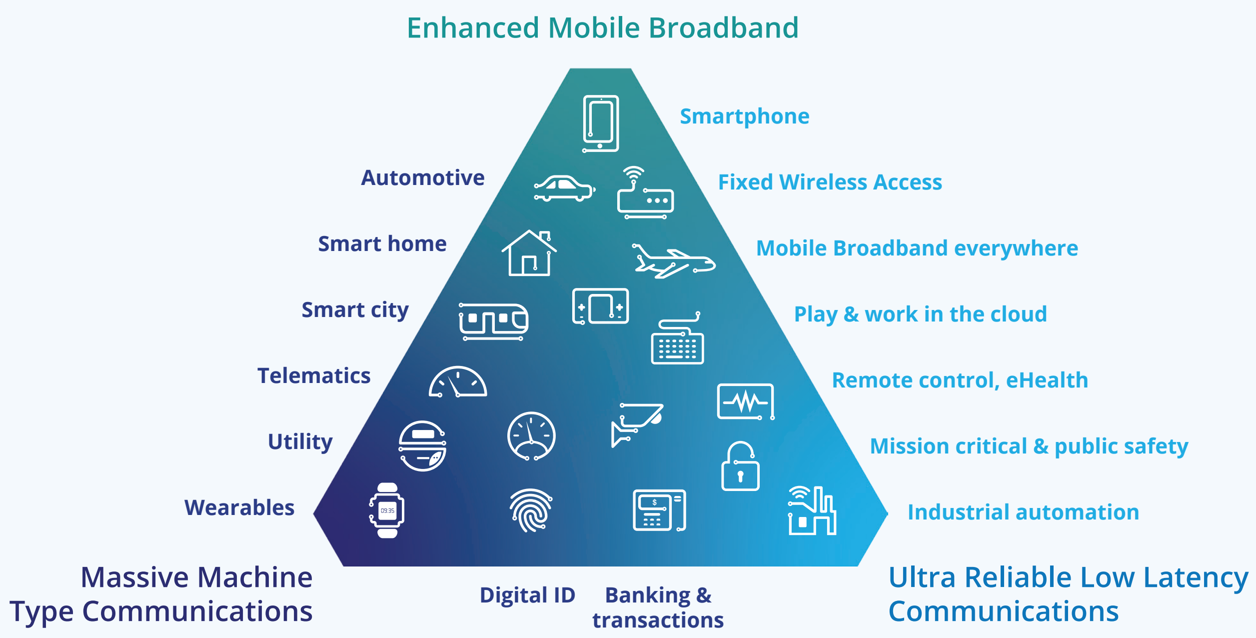
The story does not end here because 5G-NR is expected to address these applications as worldwide carriers migrate their networks to 5G, and then 6G.

3. Challenges of 5G-NR for IoT

Solutions designed based on 3GPP Release 15 and Release 16 for eMBB or urLLC do not efficiently address the needs of IoT. The industry has spent years optimizing networks, semiconductor solutions, and modules to address the needs of massive IoT with Cat 1, Cat 1bis, LTE-M and NB-IoT, in terms of scale, price, and power consumption.

To develop a dedicated alternative for the lower LTE categories, 5G began evolving, in 3GPP Release 17, to a

Figure 3: 5G-NR for IoT: eMBB, urLLC, and mMTC



device category called NR RedCap (NR Reduced Capability). NR RedCap allows for a 5G NR alternative to 4G LTE Cat 4, supporting a few 10s of Mbps (typically less than 150 Mbps). It has become clear to the industry that a solution dedicated for eMBB with more than 1Gbps speed would not be able to replace an LTE Cat 4 module priced below $25. The gap is even wider with lower LTE categories.

4. The Way Forward in IoT with 5G-NR

RedCap as an alternative is more than a mere replacement of LTE Cat 4, RedCap also brings all the benefits initiated by the 3GPP within the framework of 4G and 5G-NR, extending their use to new applications.

NR RedCap is future-proof in that it can operate fully on an optimized standalone 5G core network and does not need a 4G core to operate. Also, in comparison to the rigidly defined LTE categories, NR RedCap includes options that allow it, under the same terminology, to tackle demanding usage (up to 240 Mbps with two-layers, and up to 256 QAM in the downlink in 20 MHz of bandwidth), as well as a lower-end usage (up to 90 Mbps with one-layer, 64 QAM in downlink in 20 MHz of bandwidth). Some low power modes called MICO, Mobile Initiated Connection Only, have also been introduced with NR RedCap, allowing it, by design, to re-use what was learned from LTE in achieving longer battery life. It can also offer improved latency over LTE, and higher network capacity. IoT gateways and high-end wearables (for XR) are a perfect fit for these applications.

The technology allows for the design of an NR RedCap module that can reach the desired price point for an LTE Cat 4 module, but that is still far away from the low-teens and single digit pricing that can be achieved with LTE Cat 1bis and LTE-M.

To develop solutions that address the mid to low-end of IoT applications for smart home, smart city, telematics, and eHealth, with modules that are even more cost- and power-optimized, and closer to what Cat 1bis or LTE-M can support, we need to continue evolving standards, and optimizing semiconductors.

5. 5G-NR eRedCap and Beyond

After 3GPP Release 17, the industry will move to Release 18 and beyond for further evolutions of RedCap. As part of an ecosystem that is focused on IoT, we need to seriously consider options beyond NR RedCap.

Release 18 addresses various options for enhanced RedCap (eRedCap) and, we believe, further evolutions of eRedCap (that we will name feRedCap within the scope of this paper) will fully address what is covered today by Cat 1, Cat 1bis, LTE-M and NB-IoT, with a future-proof NR solution.

There are various ways to reach the price/performance point of these LTE categories with 5G-NR, and trade-offs can be found to optimize network capacity, coverage, and user experience.

Among them are the following:

1. **Number of receive paths**, as adopted in LTE Cat 1bis, LTE-M, and NB-IoT:
   - Decreasing them from 2 to 1 allows a typical gain at module level of 20 percent in price, and 15-30 percent in power consumption, while limiting the real estate (at module and device level), further reducing the overall solution cost.
   - To avoid negatively impacting coverage, various techniques can be implemented on the UE and/or network sides, including Release 17 coverage enhancement techniques and advanced codebook-based precoding (preferably eType II), allowing for a performance improvement of 4 to 6dB.
2. **Lower bandwidth,** as adopted in LTE-M and NB-IoT:
   - As most of the IoT applications do not require throughput exceeding a few 100s of kbps, it is obvious that we need to use channel bandwidth narrower than what has been designed for more than 100s of Mbps applications. Reducing the bandwidth of the baseband from 20 MHz (RedCap) down to 5 MHz or even down to 1 MHz has a significant impact on the cost and power consumption and can still achieve the required performance in terms of throughput.
3. **HD-FDD** (half duplex, frequency division duplex), as adopted in LTE-M and NB-IoT:
   - The majority of the 3GPP frequency bands are FD-FDD (full duplex, frequency division duplex). This means the channel bandwidth is split into two sub-bands that can be simultaneously used: one for transmission and one for reception. Consequently, the front-end radio on the device side must integrate an analog filter to isolate the receive path from the transmit path and this must be done for each of the 3GPP bands supported by the device. This has a major cost impact on the device, specifically if the device is supporting many 3GPP bands for global operation.
   - With LTE-M and NB-IoT, 3GPP came up with a very elegant way to solve such a cost issue, notably, by enabling the device to operate in half duplex HD-FDD mode that will avoid the front-end radio having to receive and transmit at the same time.[1] Consequently, HD-FDD allows a simplified implementation of the radio front-end with no need for special filtering, even for a worldwide SKU module operating on all 3GPP frequency bands. This feature was instrumental in reducing the cost of LTE-M modules down to mid-single digit pricing, and we believe allowing HD-FDD option on 5G-NR is key to bringing down the cost of 5G-NR IoT modules in the future in the same way we have done it on 4G.

Having these techniques implemented in semiconductor and module products will inevitably allow for an offering equivalent to what NB-IoT, LTE-M, Cat 1, and Cat 1bis enable today, with modules ranging from a single-digit up mid-to-high teens, supporting 5G-NR and fallback to their 4G sibling.

These various options are summarized in Table 1 on the following page.

Table 1: 5G RedCap Options

5G-NR (fe)RedCap Product Feature	5G-NR RedCap, high end	5G-NR RedCap, low end	5G-NR eRedCap	5G-NR feRedCap[2]
Baseband Bandwidth	20MHz	20MHz	5MHz	1MHz
Duplexing	FD-FDD	FD-FDD	FD or HD-FDD	HD-FDD
Number of layers	2	1	1	1
Max QAM	256	64	64	16
Max DL PHY-layer data rate	240 Mbps	90 Mbps	20 Mbps	2 Mbps
Equivalent LTE category	Cat 4	Cat 4 / Cat 1	Cat 1 / Cat 1bis	LTE-M
Price target (100% Cat 4 reference)	115%	65-75%	40-55%	30-35%

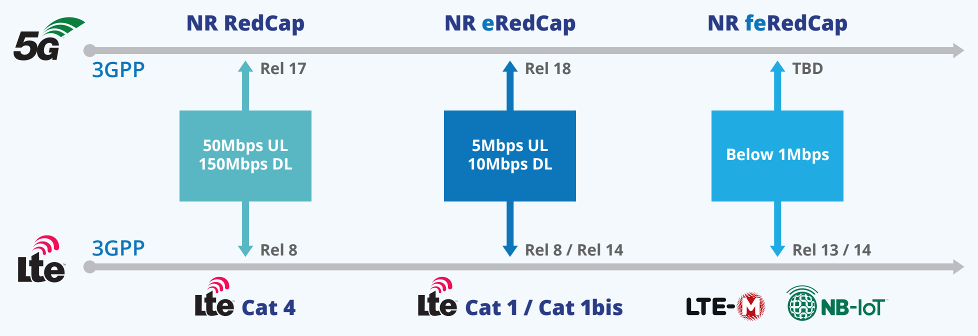
6. Conclusion

When we look beyond the existing 3GPP Release 17, the scope of the 5G-NR offering enables a true path forward to transition from 4G Cat 1 (or Cat 1bis) and LTE-M, and NB-IoT to 5G-NR technologies with similar power and cost advantages. This provides the ecosystem with a valid capacity to maintain and grow cellular IoT applications.

Three to four Reduced Capability variants of 5G-NR can be identified and will be refined by the semiconductor vendors when they will optimize their portfolio.

We strongly encourage all ecosystem actors to enable this flexible configuration and module options, offering the IoT device ecosystem a smooth transition to 5G-NR without cost and power impact on devices. Hence truly enabling a cost-efficient and sustainable IoT world.

Figure 4: 5G-NR RedCap Variants and Their 4G LTE Equivalent Categories



7. Footnotes

- [1] HD-FDD existed before LTE-M and NB-IoT but was not thought a viable solution in 3GPP until Release 13. Note that LTE-M and NB-IoT can operate in FD-FDD and in TDD, but the success of those standards relies on truly enabling an HD-FDD usage.
- [2] There is no ongoing work in 3GPP on what we qualify here as feRedCap, but we can anticipate tentative directions to fit the demand, whether part of 5G-NR or even early 6G.