5G NR Release 16: Start of the 5G Evolution

Stefan Parkvall, Yufei Blankenship, Ricardo Blasco, Erik Dahlman, Gabor Fodor, Stephen Grant, Erik Stare, and Magnus Stattin

Abstract

Standardization of the first release of 5G New Radio has been completed in the 3rd Generation Partnership Project, and the evolution has started. On a high level, the enhancements in Release 16 can be grouped into two categories: improvements of existing features, and introduction of new features addressing new verticals and deployment scenarios. This article provides a brief overview of the enhancements in the first category, followed by a more in-depth discussion of features in the second category. These features include integrated access-backhaul, operation in unlicensed spectrum, intelligent transportation and vehicular-to-anything communication, the Industrial Internet of Things, ultra-reliable low-latency communication, and positioning. The article concludes with an outlook on enhancements planned for Release 17.

Introduction

With 4G cellular systems in service for almost 10 years, we are now seeing the introduction of 5G with several networks already in commercial operation. The highly flexible 5G New Radio (NR) radio access technology specified by the 3rd Generation Partnership Project (3GPP) is already, from its first incarnation (Release 15), capable of supporting a wide range of requirements including extremely high data rates and very low latency.

NR uses orthogonal frequency-division multiplex (OFDM) and operates in frequency bands up to 52.6 GHz using either time-division duplex (TDD) or frequency-division duplex (FDD). Dynamic TDD is supported where the transmission direction (uplink or downlink) can be dynamically controlled by the base station (BS). NR can operate in either a non-standalone mode, where the 4G LTE system is used for mobility and initial access, or standalone, without dependence on another radio access technology. Key NR features include extensive multi-antenna support, including spatial multiplexing as well as beamforming for both data transmission and initial access. Ultra-lean transmission, that is, minimizing transmission of "always-on" signals present regardless of the user traffic, is a key design principle for NR. An important aim in the development of the NR specification has been to ensure a high degree of forward compatibility in the radio interface design. In this context, forward compatibility implies a design that allows for substantial future evolution, in terms of introducing new technology and enabling new services with as yet unknown requirements and characteristics. At the same time, legacy user equipments (UEs, the 3GPP name for terminals) on the same carrier are supported. A detailed description of the NR technology is beyond the scope of this article, and the reader is referred to [1–4] for details on Release 15.

With the ongoing work on Release 16, 3GPP has already started the evolution of NR. Although Release 16 is a significant enhancement to the NR standard, it is expected that the evolution will continue for many years with Release 16 merely being the first of several steps of the NR evolution. In this article, we discuss some of the major enhancements brought by Release 16 focusing on the physical layer, as well as outline what will happen in Release 17.

NR Release 16

As mentioned, NR Release 16 represents a significant evolution of NR and contains several key enhancements. On a high level, the enhancements can be grouped into two categories: improvements of already existing features, and new features addressing new deployment scenarios and verticals.

In the first category, enhancements in Release 16 include:

- Cross-link interference (CLI) mitigation and remote interference management (RIM), which address two interference scenarios in TDD systems [5]. CLI mitigation primarily targets small cell deployments using dynamic TDD and introduces new measurements to detect crosslink interference between downlink and uplink to help to schedule transmissions. RIM, on the other hand, targets wide-area TDD deployments where certain weather conditions can create atmospheric ducts leading to interference from very distant BSs. With RIM, problematic interference scenarios can be managed in an automated way in contrast to today's largely manual intervention approaches.
- Multiple-input multiple-output (MIMO) enhancements [6], which introduce enhanced beam handling and channel state information feedback, as well as support for transmission to a single UE from multiple transmission points (multi-TRP). These enhancements increase throughput, reduce overhead, and/or provide additional robustness.
- UE power savings [7], which introduce features such as a wake-up signal and enhancements to control signaling and scheduling mechanisms to reduce the UE power consumption.

Digital Object Identifier:
10.1109/MCOMSTD.011.1900018

The authors are with Ericsson Research.

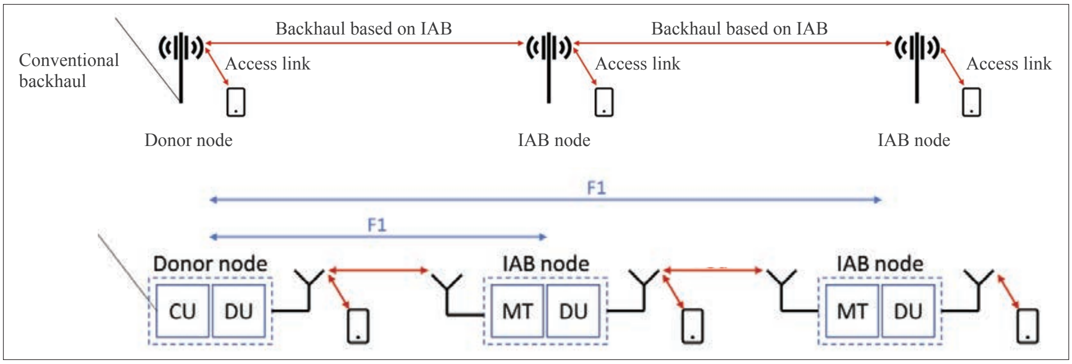

FIGURE 1. High-level architecture of IAB.

- Dual connectivity (DC) and carrier aggregation (CA) enhancements [8], which offer reduced latency for setup and activation of CA/DC, thereby leading to improved system capacity and data rates. For example, in Release 15, a UE connection can be resumed after periods of inactivity without the need for extensive signaling. Measurement configuration and reporting, however, does not take place until the UE enters the fully connected state. Enhancements in Release 16 will enable reporting of measurements and restoration of saved CA/DC configurations as part of the resume procedure to allow faster activation of CA/DC with less signaling.
- Mobility enhancements [9], which offer reduced handover delays, in particular when applied to beam management mechanisms used for deployments in millimeter-wave (mmWave) bands.

One should not underestimate the benefits of gradual improvements such as those listed above, since they indeed contribute to making NR an even more efficient system. However, due to space limitations, the remainder of the article primarily focuses on features targeting new verticals and/or new deployment scenarios.

Integrated Access and Backhaul

Integrated access and backhaul (IAB) extends NR to also support wireless backhaul as an alternative to, for example, fiber backhaul [10]. This enables the use of NR for a wireless link from central locations to distributed cell sites and between cell sites. For example, this can simplify deployments of small cells in dense urban networks or enable deployment of temporary sites used for special events.

IAB can be used in any frequency band in which NR can operate. However, it is anticipated that mmWave spectrum will be the most relevant spectrum for IAB due to the amount of spectrum available. Higher-frequency spectrum also enables a higher degree of beamforming at both the transmitter and receiver, something that is especially suitable for the stationary conditions of a backhaul link and will reduce the overall interference level. As higher-frequency spectrum is typically unpaired, this also means that IAB can be expected to primarily operate in TDD mode on the backhaul link. Furthermore, the backhaul link may either operate in the same frequency band as the access link, known as *inband* operation, or use a separate frequency band (*out-of-band* operation).

Overall Architecture

Figure 1 illustrates the basic structure of a network utilizing IAB. An *IAB node* connects to the network via an (IAB) *donor node*, which essentially is a normal BS utilizing conventional (non-IAB) backhaul. The IAB node creates cells of its own and appears as a normal BS to UEs connecting to it. Thus, there is no specific UE impact from IAB, which is solely a network feature. This is important as it also allows for legacy (Release 15) UEs to access the network via an IAB node. As illustrated in Fig. 1, additional IAB nodes can connect to the network via the cells created by an IAB node, thereby enabling multihop wireless backhauling.

Architecture-wise, IAB is based on the so-called *CU/DU split* already introduced in Release 15. The CU/DU split implies that, specification-wise, the BS is split into two parts: a centralized unit (CU) and one or more distributed units (DUs) where the CU and DU(s) may be physically separated depending on the deployment. The CU includes higher-layer protocols, that is, radio resource control and packet data convergence protocol. The DU includes the lower-layer protocols, that is, radio-link control (RLC), medium access control (MAC), and the physical layer. The CU and DU are connected via a standardized F1 interface.

In principle, IAB is an NR-based wireless link for the F1 interface. As illustrated in the lower part of Fig. 1, an IAB node thus includes a conventional DU part creating cells to which UEs and other IAB nodes can connect. The IAB node also includes a mobile termination (MT) part providing connectivity for the IAB node to (the DU of) the donor node. In the case of multihop IAB, the MT part provides connectivity to the DU part of higher-level IAB nodes. This parent-DU-to-MT link separately implements the per-link physical, MAC, and RLC protocols.

In many respects, the MT functionality is similar to a UE. In fact, when initially accessing the network, the IAB-node MT does that in the same way as a UE. Once the MT has connected to a parent node and, eventually, to a donor node, the DU of the IAB node is configured and activated.

In many respects, the MT functionality is similar to a UE. In fact, when initially accessing the network, the IAB-node MT does that in the same way as a UE. Once the MT has connected to a parent node and, eventually, to a donor node, the DU of the IAB node is configured and activated.

Physical-Layer Impact

On the physical layer, the main extension of IAB is to enable in-band relaying, that is, the same frequency band is used for the MT and DU side of

In Release 16, NR is enhanced to enable operation in unlicensed spectrum. LTE was extended to unlicensed spectrum in Release 13, and this is known as licensed assisted access (LAA), where the UE is attached to the network using a licensed carrier, and one or more unlicensed carriers are used to boost data rate using the carrier aggregation framework.

the IAB node. In practice, the DU and MT of an IAB node typically cannot operate simultaneously, since the transmission at one side would cause severe interference to the reception on the other side (*half-duplex* constraint). This is addressed by configuring DU time-domain resources for the access link as *hard* or *soft*.

A hard DU resource can always be used by the IAB node without considering the impact on its ability to transmit or receive on the MT side. In contrast, a soft DU resource can only be used conditionally, either by the IAB node concluding it can use the resource without impacting the MT (implicit indication of availability), or by the IAB node being explicitly informed by the parent node about the instantaneous availability of the resource (explicit indication). The possibility for soft DU resources, in combination with explicit and implicit indication of availability, allows for more dynamic sharing of time-domain resources between the DU and the MT part of the IAB node.

NR in Unlicensed Spectrum

Spectrum availability is essential to wireless communication, and the large amount of spectrum available in unlicensed bands is attractive for increasing data rates and capacity for 3GPP systems. In Release 16, NR is enhanced to enable operation in unlicensed spectrum. LTE was extended to unlicensed spectrum in Release 13, and this is known as licensed assisted access (LAA), where the UE is attached to the network using a licensed carrier, and one or more unlicensed carriers are used to boost data rate using the carrier aggregation framework. NR supports a similar setup; however, unlike LTE, standalone operation in unlicensed spectrum is supported, that is, operation without assistance from a carrier in licensed spectrum. This is one of the main new features compared to LTE-LAA and will greatly add to the deployment flexibility.

The aim in Release 16 standardization was to have a single global framework that allows operation not only in the existing 5 GHz unlicensed bands (5150–5925 MHz), but also in new bands when they become available. The most important new band is the 6 GHz band (5925–7125 MHz), partially or fully available in large parts of the world, where there are potentially hundreds of megahertz available. While both NR and WiFi will be new to the 6 GHz band, it is assumed that regulations will provide a framework for protection of incumbent services (e.g., fixed services) when unlicensed technologies operate in this band.

Several key principles important for operation in unlicensed spectrum are already part of NR in Release 15. For example, ultra-lean transmission, that is, transmissions only occur when there is data to deliver to a UE, is essential to operation in unlicensed spectrum. Another example is the NR frame structure, which was designed such that transmissions can start at any OFDM symbol and not only at the slot boundary, sometimes referred to as "mini-slot transmission." This is beneficial for operation in unlicensed spectrum as it allows the system to quickly access the channel once it is available. This benefit is enhanced further by the use of larger subcarrier spacings (30 and 60 kHz) compared to LTE-LAA (15 kHz), resulting in shorter OFDM symbol durations. Nevertheless, there are still some extensions needed in NR to enable operation in unlicensed spectrum, and such extensions are the focus of the Release 16 standardization effort.

Listen Before Talk

Channel access is probably the most obvious area of enhancement [11]. Prior to accessing the channel for any downlink or uplink transmission, the initiating node (BS or UE) randomly initializes a backoff counter between 0 and an initial contention window size that depends on the priority of the traffic being served. A clear-channel assessment (CCA) check is performed in a series of 9 μs CCA slots. For each successful CCA, the backoff counter is decremented until it reaches zero, upon which a transmission can occur. The amount of time that the initiating node may occupy the channel is referred to as the channel occupancy time (COT), and this depends on the priority class for the served traffic. Typical values are in the range of 2–10 ms. If a collision on the channel occurs (indicated by *hybrid automatic repeat request*, HARQ, feedback), the contention window size is doubled for the next channel access by that node. Successful transmissions allow the contention window size to be reset to its initial value. Such a mechanism is often referred to as listen before talk (LBT).

NR largely reuses the same LBT mechanism as for LAA with some enhancements to support multiple downlink/uplink switching points within a COT, and support of LBT with the appropriate channel access priority class for signals that were not subject to LBT in LTE-LAA where, for example, system information reception and random access is handled using the licensed carrier. Reusing the mechanism developed for LTE-LAA (based largely on the mechanism used by WiFi) is beneficial from a coexistence perspective in the 5 GHz band. In fact, it was demonstrated during the studies in 3GPP that replacement of one WiFi network with an NR network can lead to improved performance for the remaining WiFi networks [11]. In the 6 GHz band, considered as greenfield spectrum, use of two detection thresholds as used by WiFi does not make sense due to lack of deployed WiFi. In this scenario, it was demonstrated in several contributions that performance is maximized for coexisting systems if all technologies use a common energy detection (ED) threshold when performing LBT, rather than different thresholds for different unlicensed technologies as is used in the 5 GHz band.

Physical Layer Enhancements

The transmission of a burst of several existing Release 15 signals/channels, specifically, the synchronization signal block used for synchronization and delivery of critical system information, the downlink data channel carrying remaining system information, the associated control channel scheduling the data, and potentially channel state information (CSI) reference signals, is referred to in Release 16 as a *discovery* burst (DB). In the case of a beamformed system, a DB can consist of multiple instances of these signals/channels, each beamformed separately. For operation in unlicensed spectrum, the components of the DB

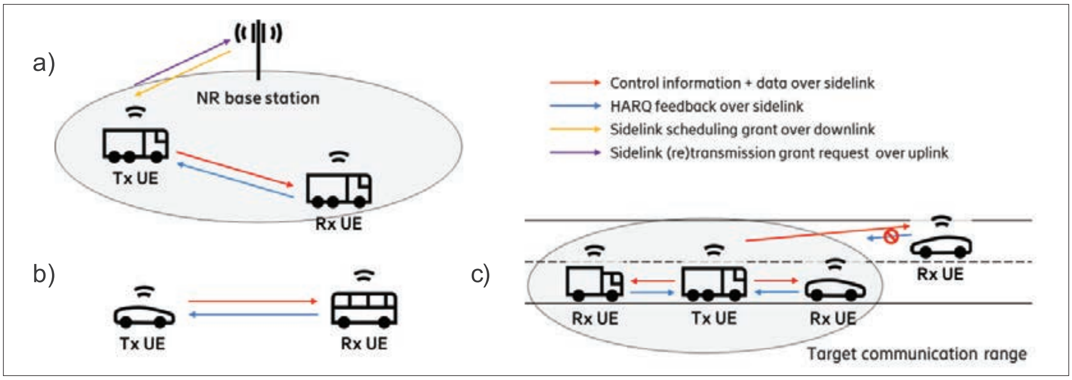

Figure 2. a) Unicast vehicle-to-vehicle (V2V) communications in cellular coverage with network control (Mode 1); b) unicast V2V communications outside of cellular coverage with UE autonomous scheduling (Mode 2); c) groupcast V2V communications with distance-based HARQ feedback.

are transmitted without time gaps in one or more consecutive slots in order to avoid separate LBT operation that would otherwise be required. Furthermore, to ensure that DB transmission is robust to LBT failures, a window is defined within which the DB can shift in time relative to the frame boundary if LBT is unsuccessful. This is important as the DB is used for fundamental functions such as initial access, idle/connected mode mobility, and time/frequency synchronization, all of which are critical to enable standalone operation.

Enhancements are also made to the uplink channels to spread the signal over a larger bandwidth to meet requirements on the maximum power spectral density and the requirements on the minimum bandwidth occupancy in some regulatory regions. This applies to the uplink data and control channels using a set of interlaced resource blocks and to the physical random access channel using a set of contiguous resource blocks spanning nearly 20 MHz. Wideband transmission of the sounding reference signal is already supported in Release 15. HARQ mechanisms are also enhanced to account for the fact that feedback may need to be delayed to the next available COT. These mechanisms also include the ability for the BS to request (re)-transmission of HARQ feedback, for example, due to LBT failure at the UE or misdetection of the uplink control channel carrying the feedback.

Intelligent Transportation Systems and Vehicle-to-Anything Communications

Intelligent transportation systems (ITS) are one example of a new vertical in focus for NR Release 16. ITS provides a range of transport and traffic management services that together will improve traffic safety and reduce traffic congestion, fuel consumption, and environmental impacts. To facilitate ITS, communication is required not only between vehicles and the fixed infrastructure but also between vehicles. To this end, 3GPP defines 25 use cases for advanced vehicle-to-anything (V2X) communications with requirements that go far beyond those of broadcasting safety messages. These advanced use cases are categorized as vehicles platooning, cooperative communication using extended sensors, advanced driving (which includes collision avoidance and cooperative lane change), and remote driving.

Accordingly, V2X is commonly used to refer to all kinds of communication to and from vehicles. Communication with the fixed infrastructure is obviously already catered for in NR Release 15 via the access link interface (Uu) between the BS and UE, while the NR sidelink (PC5) is added in Release 16. Besides the sidelink, many of the enhancements introduced for the cellular uplink/downlink interface of NR in Release 16 are also relevant for supporting ITS services. In particular, ultra-reliable low-latency communication (URLLC) is instrumental in enabling remote driving services.

Sidelink Design

To support a wide range of services including vehicle platooning, extended sensors, and advanced driving, the NR Release 16 sidelink is equipped with a broad set of new physical and MAC layer capabilities [12]. Specifically, the sidelink enables physical-layer unicast, groupcast, and broadcast communication, thereby providing highly reliable communication links with appropriate quality of service (QoS) characteristics as demanded by ITS services. The sidelink can operate in in-coverage, out-of-coverage, and partial coverage scenarios (Fig. 2), and can make use of all NR frequency bands.

Supporting unicast sidelink communications (Figs. 2a and 2b) is a fundamental new capability of the NR Release 16 sidelink, compared to LTE, and is the foundation of providing a highly reliable and QoS-enabled physical layer for ITS services. Indeed, the unicast-capable physical layer of the sidelink allows a transmitter node to acquire CSI with respect to the intended receiver, adapt the transmission characteristics to the prevailing channel conditions, and employ HARQ and link adaptation to enhance reliability. For this purpose, a sidelink CSI acquisition framework has been introduced in Release 16. Similar to the Uu downlink, the use of CSI reference signals allows the receiver to estimate and report the channel quality and, in the case of multi-antenna transmissions, the recommended transmission rank. However, there are important differences compared to downlink CSI acquisition as reporting a precoder matrix is not supported in sidelink, and the CSI reference signals are conveyed in-band with a sidelink transmission, limiting the possibili-

To support a wide range of services including vehicle platooning, extended sensors, and advanced driving, the NR Release 16 sidelink is equipped with a broad set of new physical and MAC layer capabilities. Specifically, the sidelink enables physical-layer unicast, groupcast, and broadcast communication, thereby providing highly reliable communication links with appropriate QoS characteristics as demanded by ITS services.

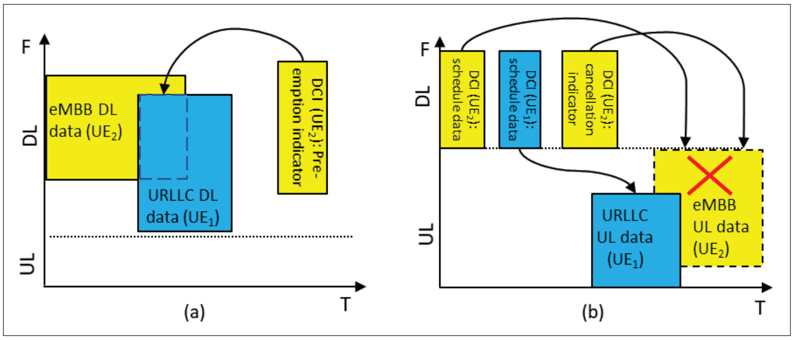

FIGURE 3. Example of a) Release 15 downlink preemption; b) Release 16 uplink cancellation.

ties to perform channel sounding. Besides channel quality and the recommended transmission rank, long-term reference signal received power measurements measured on demodulation reference signals are fed back to the transmitter. The power measurements are used for power control, helping to maintain a predefined signal-to-noise ratio target and thereby a service-specific quality of the sidelink.

Building on the extended physical and MAC layer capabilities, the NR sidelink adopts the concept of QoS-enabled radio bearers that serve as the access stratum foundation on which the higher layers of ITS services can be created. These new features of the NR sidelink make it resemble the NR uplink in terms of adapting UE transmission parameters and maintaining QoS targets subject to energy efficiency, battery, and regulatory constraints.

In addition to unicast, groupcast communications also play an important role for V2X. Two ways of operating groupcast communications are introduced in Release 16. The first one, which targets groups that are stable over relatively long periods of time (e.g., truck platoons), enables high reliability by means of group establishment and dedicated HARQ signaling between each of the receivers and the transmitter. In contrast, the second way of operating groupcast communications (Fig. 2c) targets dynamic groups that are defined in terms of distance between transmitter and receiver(s). For example, a cooperative lane-merging message is typically relevant for all UEs within a certain distance from the transmitter. For these cases, the NR sidelink supports groupcast communications where the relevance area of a message determines the UEs that are involved in the HARQ retransmission. That is, HARQ retransmissions are only requested by UEs that are within the relevance area, which reduces overhead and interference levels. This approach makes transmitter-receiver distance and desired communication range central to the notion of QoS in sidelink groupcast communications for V2X.

To support a rich set of ITS services — beyond cooperative-awareness messages and decentralized environmental notification messages — with differences in periodic or aperiodic packet arrival, different arrival rates, and differences in latency and error rate requirements, the NR sidelink provides flexible resource allocation procedures with varying degrees of network control. Mode 1 (Fig. 2a), which is used for operation under network control, allows for tight UE control by the BS via dynamically scheduling every sidelink transmission as well as for semi-static allocations, on which the UEs largely manage their own sidelink transmissions. When the sidelink operates without network assistance, in Mode 2 (Fig. 2b), the transmitting UEs employ sensing procedures for collision avoidance and use resource selection and reservation procedures that facilitate high resource utilization and low packet error rates.

Industrial IoT and Ultra-Reliable Low-Latency Communication

The Industrial Internet of Things (IIoT) is another major vertical in focus for NR Release 16. While Release 15 can provide very low air interface latency and high reliability [13], further enhancements to latency and reliability are introduced in Release 16. This is to enable widening of the set of industrial IoT use cases and to include increased demand for new use cases, such as factory automation, electrical power distribution, and the transport industry (including the remote driving use case). Time-sensitive networking (TSN), where latency variations are as important as low average latency, is one area addressed with introduction of RRC signaling providing accurate time synchronization of the UEs. Another example is a mechanism to prioritize traffic flows within and between UEs. In general, many of the additions can be viewed as a collection of smaller improvements, some of which are discussed below, that together significantly enhance NR in the area of URLLC [14].

Uplink Preemption

A later scheduled low-latency higher-priority communication of UE1 may need to preempt a previously scheduled (lower-priority) transmission of UE2, that is, to allow immediate UE1 transmission by cancelling a part of or the whole originally scheduled UE2 transmission. Inter-UE downlink preemption is already supported in Release 15 (Fig. 3a), where the BS can always decide to schedule another transmission to UE1 using the resources allocated to a previously scheduled transmission to UE2, and UE2 is subsequently notified and can discard this transmission. In Release 16, this is extended to the uplink such that a previously scheduled lower-priority uplink transmission of UE2 can be preempted (i.e., cancelled) by a higher-priority uplink transmission of UE1 (Fig. 3b). In the downlink, a UE is notified of the preemption after the preemption has occurred, which enables the UE to discard the destroyed soft bits from its soft buffer. In contrast, for the uplink, the UE is notified of the preemption before the previously scheduled transmission happens. This implies that UE2 scheduled with lower-priority uplink transmission should monitor the cancellation indicator, part of the downlink control information, frequently before actual transmission if the BS notifies it to cancel the uplink transmission. When notified, the affected UE2 cancels its uplink transmission.

In addition to cancellation of ongoing lower-priority uplink transmissions, enhanced power control to increase the transmission power of a higher-priority uplink transmission is also part of Release 16.

Uplink Collision Resolution

Handling of intra-UE uplink resource conflicts is also supported in Release 16. Examples where multiple uplink channels and traffic flows compete for the same resources are the following: between dynamic grant and configured grant transmissions, conflicts involving multiple uplink configured grants, uplink control vs. data collisions, and uplink control vs. control collisions. Such conflict handling will primarily improve the resource efficiency when mixing different traffic flows in the same system, thereby enabling smooth introduction of URLLC into cellular networks. The collision of uplink transmission is resolved in two steps: first, the collision among uplink transmissions of the same priority are resolved, followed by resolving collisions between uplink transmissions with different priority by dropping the lower priority transmission(s).

Accordingly, each type of uplink control (e.g., scheduling requests, HARQ Acknowledgments, CSI feedback) and data is assigned a priority for collision resolution. For example, priority of dynamically scheduled uplink data is indicated by a priority indicator field in the scheduling grant, while the priority of uplink configured grant data is provided via the RRC configuration.

Downlink Control Channels

NR is fundamentally a scheduled system where each UE monitors a set of downlink control channels to determine whether it is scheduled to transmit or receive. To reduce latency, Release 16 supports more frequent control channel monitoring, up to every second OFDM symbol in the extreme case (compared to once per slot, as in Release 15). To ensure the monitoring burden is evenly spread among monitoring spans, Release 16 specifies the amount of downlink control channel monitoring in each monitoring span. This alleviates the UE implementation complexity when compared to Release 15, where the demand on UE monitoring is defined over one slot.

To ensure the monitoring burden is evenly spread among monitoring spans, Release 16 specifies the amount of downlink control channel monitoring in each monitoring span. This alleviates the UE implementation complexity when compared to Release 15, where the demand on UE monitoring is defined over one slot.

Configured Grants and Semi-Persistent Scheduling

Uplink-configured grant transmission is a key mechanism to enable low-latency data transmission by pre-allocating resources to avoid the scheduling request/scheduling grant phase prior to uplink data transmission. In the downlink, semi-persistent scheduling (SPS) is a useful tool to reduce control signaling. In Release 16, for both uplink configured grant and downlink semi-persistent scheduling, multiple configurations can be active simultaneously to support multiple services, where each service may have a different performance requirement in terms of latency and reliability. For SPS, Release 16 supports a reduction in the shortest SPS periodicity down to 1 slot (e.g., 0.125 ms if the subcarrier spacing is 120 kHz). This is much shorter than the corresponding value of 10 ms in Release 15. The enhanced downlink semi-persistent scheduling and uplink configured grants are especially useful for TSN traffic, where the traffic pattern (e.g., message periodicity, message arrival time) are known to the BS.

Positioning

Global navigation satellite systems (GNSSs), assisted by cellular networks, have for many years been used for UE positioning. This provides accurate positioning but is typically limited to outdoor areas with satellite visibility. Currently, there is a range of applications requiring accurate positioning not only outdoors but also indoors. For example, the U.S. Federal Communications Commission requires 50 m horizontal accuracy and floor-level vertical accuracy for positioning of emergency calls. Industrial applications, such as manufacturing and logistics, may sometimes require very accurate positioning. Additional positioning methods beyond GNSS are therefore of great importance, especially for indoor applications, and Release 16 adds enhancements in this area.

Architecture-wise, NR positioning is based on the use of a location server, similar to LTE. The location server collects and distributes information related to positioning (UE capabilities, assistance data, measurements, position estimates, etc.) to the other entities involved in the positioning procedures (BSs and UEs). A range of positioning methods, both downlink-based and uplink-based, are used separately or in combination to meet the accuracy requirements for different scenarios.

Downlink-Based Positioning

Downlink-based positioning is supported by providing a new reference signal called the positioning reference signal (PRS). Compared to LTE, the PRS has a more regular structure and much larger bandwidth, which allows for more precise correlation and time of arrival (TOA) estimation. The PRS is periodically transmitted by each cell with the periodicity configurable from a few milliseconds up to 10 s. Since the UE may need to detect PRS from distant cells, a critical aspect of downlink PRS design is to combat interference. This is achieved by defining a number of orthogonal signals where different PRSs are separated using different subcarriers and, additionally, are separated in the time domain. The UE can then report the TOA difference for PRSs received from multiple distinct BSs, and the reports are used by the location server to determine the position of the UE. If different PRSs are transmitted in different beams, the reports will indirectly give information in which direction from a BS the UE is located.

Uplink-Based Positioning

Uplink-based positioning is based on Release 15 sounding reference signals (SRSs) with Release 16 extensions. Hearability at non-serving cells is improved by increased SRS length, open-loop power control, and beamforming procedures. A larger comb size and staggered comb patterns are also used to allow for increased range in combination with orthogonal SRSs. Based on the received SRSs, the BSs can measure and report (to the location server) the arrival time, the received power, and the angle of arrival.

The time difference between downlink reception and uplink transmission can also be reported

Area	Brief description
RedCap (NR light)	Reduced-capability NR, targeting mid-tier applications such as machine-type communication (MTC) for industrial sensors, video surveillance, and wearables. Data rates targeted are between NB-IoT/LTE-M data rates and "full" NR data rates. Extended battery lifetime is important at least for some applications.
NR above 52.6 GHz	Extending NR to the frequency range 52.6–71 GHz to allow even more spectrum to be exploited, including the 60 GHz unlicensed band.
Multicast/broadcast	Introduction of broadcast/multicast capabilities to NR, primarily targeting V2X, public safety, IP multicast, software delivery, and IoT applications.
Small data transmission	Enhancements targeting transmission of small data packets in inactive state to reduce the overhead from connection establishment. Use cases include keep-alive messages, wearables, and various sensors.
Non-terrestrial networks	Introducing support for satellites and high-altitude platforms as an additional mean to provide coverage in rural areas.
XR	Evaluation of the needs in terms of simultaneously providing very high data rates and low latency in a resource-efficient manner to support various forms of augmented reality (AR) and virtual reality (VR), collectively referred to as XR.
IAB enhancements	Enhancements to Release 16 IAB to support (limited) network topology changes and improved duplexing of access and backhaul links.
Sidelink enhancements	Enhancements to the Release 16 sidelink to improve resource efficiency and reliability, as well as studying sidelink relaying.
Industrial IoT enhancements	Improved support for factory automation and URLLC, including synchronization enhancements for time-sensitive networking.
Coverage enhancements	Study possibilities to provide enhanced (wide-area) coverage.
Positioning enhancements	Enhancements to Release 16 positioning mechanism, in particular sub-meter accuracy for industrial use cases.
RAN data collection enhancements	Improved mechanisms for collecting measurement information from UEs and network in order to simplify deployment and enhance the support self-optimized networks.
MIMO enhancements	Improvements to Release 16 MIMO support based on experience from commercial networks.
Power-saving enhancements	Enhanced mechanisms for UE power saving, for example improvements in the area of discontinuous reception and blind decoding of control channels.
Dynamic spectrum sharing enhancements	Dynamic spectrum sharing, facilitating NR and LTE to share the same carrier, is supported already in Release 15 but will be enhanced with e.g. cross-carrier scheduling and other scheduling enhancements.
Dual-connectivity enhancements	Continuation of Release 16 dual connectivity enhancements, for example improvements to activation/deactivation of secondary cells.

Table 1. Candidate enhancements discussed for possible inclusion in Release 17.

and used in round-trip time (RTT)-based positioning schemes where the distance between a BS and a UE can be determined based on the estimated RTT. By combining several such RTT measurements, involving different BSs, the position can be determined.

Release 17 Outlook

Looking beyond Release 16, NR will continue to evolve in Release 17. The initial content of Release 17 was agreed in December 2019, most of which is listed in Table 1. Some of the potential enhancements (e.g., MIMO enhancements and IAB enhancements) are improvements to features present in Release 16, while others (e.g., RedCap, NR beyond 52.6 GHz, and non-terrestrial networks) address new scenarios not supported in Release 16.

Conclusions

From the discussion above, it is clear that NR is rapidly evolving, starting in 3GPP Release 16. The emphasis on forward compatibility in the basic NR design ensures that the introduction of extensions in most cases is relatively straightforward. NR is a very flexible platform capable of evolving in a wide range of directions and an attractive path to future wireless communication.

References

[1] E. Dahlman, S. Parkvall, and J. Sköld, *5G NR — The Next Generation Wireless Access Technology*, Academic Press, Oxford, U.K., 2018.

[2] S. Parkvall *et al.*, "NR: The New 5G Radio Access Technology," *IEEE Commun. Standards Mag.*, Dec. 2017.

[3] H. Ji *et al.*, "Ultra-Reliable and Low-Latency Communications in 5G Downlink: Physical Layer Aspects," *IEEE Wireless Commun.*, June 2018, pp. 124–30.

[4] A. Ghosh *et al.*, "5G Evolution: A View on 5G Cellular Technology Beyond 3GPP Release 15," *IEEE Access*, vol. 7, Sept. 2019, pp. 127639–51.

[5] 3GPP TR 38.828, "Cross Link Interference (CLI) Handling and Remote Interference Management (RIM) for NR."

[6] 3GPP RP-182863, "Enhancements on MIMO for NR."

[7] 3GPP TR 38.840, "Study on User Equipment (UE) Power Saving in NR."

[8] 3GPP RP-191600, "LTE-NR & NR-NR Dual Connectivity and NR Carrier Aggregation Enhancements."

[9] 3GPP RP-190489, "NR Mobility Enhancements."

[10] 3GPP TR 38.874, "NR; Study on Integrated Access and Backhaul."

[11] 3GPP TR 38.889, "Study on NR-Based Access to Unlicensed Spectrum."

[12] 3GPP TR 38.885, "Study on NR Vehicle-to-Everything (V2X)."

[13] 3GPP TR 38.913, "Study on Scenarios and Requirements for Next Generation Access Technologies."

[14] 3GPP TR 38.824, "Study on Physical Layer Enhancements for NR Ultra-Reliable and Low Latency Case (URLLC)."

[15] 3GPP TR 38.825, "Study on NR Industrial Internet of Things (IoT)."

[16] 3GPP TR 38.855, "Study on NR Positioning Support."

Biographies

Stefan Parkvall [F] joined Ericsson in 1999 and is a senior expert working on 6G and future radio access. He is one of the key persons in the development of HSPA, LTE, and NR radio access, and has been deeply involved in 3GPP standardization for many years. He is a co-author of several popular books such as *4G — LTE/LTE-Advanced for Mobile Broadband* and *5G NR — The Next Generation Wireless Access*. He has more than 1500 patents in the area of mobile communication and holds a Ph.D. in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden.

Yufei Blankenship received her Ph.D. degree in electrical engineering from Virginia Tech in 2000. Since then she has been actively participating in physical layer standardization with experience in 3GPP NR/LTE/HSPA and IEEE WiMAX/WLAN. Currently, she is a standards researcher with Ericsson. For 3GPP NR Releases 15 and 16, her focus has been on channel coding and URLLC. She holds numerous patents and publications in the field of wireless communications.

Ricardo Blasco holds M.Sc. (2007) and Ph.D. (2013) degrees in telecommunications from the Technical University of Catalonia (UPC), Spain, and KTH, respectively. Since 2014 he has been with Ericsson Research, working on device-to-device and vehicular communications in 4G and 5G networks, and as a standardization delegate in 3GPP RAN WG1.

Erik Dahlman joined Ericsson in 1993 and is currently a senior expert in Radio Access Technologies within Ericsson Research. He has been involved in the development of wireless access technologies from early 3G, via 4G LTE, to, most recently, 5G NR. He is currently focusing on the evolution of 5G as well as technologies applicable to beyond 5G wireless access. He is the co-author of the books *3G Evolution — HSPA and LTE for Mobile Broadband*, *4G — LTE and LTE-Advanced for Mobile Broadband*, *4G — LTE-Advanced Pro and the Road to 5G*, and *5G NR — The Next Generation Wireless Access Technology*. He has a Ph.D. in telecommunication from KTH.

Gabor Fodor [SM'08] received his Ph.D. degree in electrical engineering from the Budapest University of Technology and Economics in 1998, his Docent degree from KTH in 2019, and his D.Sc. degree from the Hungarian Academy of Sciences in 2019. He is currently a master researcher with Ericsson Research and an adjunct professor with KTH. He was a co-recipient of the IEEE Communications Society Stephen O. Rice Prize in 2018. He is serving as the Chair of the IEEE Communications Society Full Duplex Emerging Technologies Initiative and as an Editor for *IEEE Transactions on Wireless Communications* and *IEEE Wireless Communications*.

Stephen Grant received his M.A.Sc. and Ph.D. degrees in electrical engineering from Simon Fraser University in Burnaby, Canada, and his B.A.Sc. degree in electrical engineering from the University of British Columbia, Vancouver, Canada, in 1996, 2000, and 1993 respectively. He joined Ericsson Research in 2001 working with MIMO, multi-user detection, and interference cancellation technologies. Since 2016 he has worked on 3GPP 5G NR standardization in the areas of mmWave communications, beam management, and unlicensed operation.

Erik Stare holds an M.Sc. degree (1984) in telecommunication from KTH. Between 1987–1992 (at Swedish Telecom) and 1992–2018 (at Teracom), he was deeply involved in developing systems and standards for digital terrestrial broadcasting, mainly in the context of DVB, WorldDAB, and ATSC. Since 2018 he has been with Ericsson Research holding a master researcher position, working on positioning in 5G networks, and since 2019 as a team leader for 5G/NR multicast/broadcast standardization and a 3GPP RAN1 delegate.

Magnus Stattin graduated and received his Ph.D. degree in radio communication systems from KTH in 2005. He joined Ericsson Research in Stockholm, Sweden, in June 2005. At Ericsson Research he has been working on research in the areas of radio resource management and radio protocols of various wireless technologies. He is active in concept development and 3GPP standardization of LTE, NB-IoT, NR, and future wireless technologies.