NOKIA

5G-Advanced: Expanding 5G for the connected world

White paper

5G-Advanced is set to evolve the 5G system to its fullest capabilities. It will comprise a large set of innovations offering a plethora of benefits for network and system operators, end-users and verticals. 5G-Advanced will include features that will be specified in 3GPP Release 18 and beyond, including, among others, improved coverage and capacity, enhanced end-user experience and expanded capabilities beyond connectivity. It will introduce enhancements needed for more demanding applications, such as mobile extended reality and haptic applications. 5G-Advanced will inject more intelligence into the network, utilizing machine learning to adapt to the environment, boost performance and manage complex optimizations. It will also bring mobile broadband to new classes of devices (e.g. extended reality, reduced capability devices, new sidelink innovations, reduced bandwidth operation) and open 5G to new sectors. Energy efficiency will be central to 5G-Advanced, both in maximizing device battery life and in reducing network power consumption. As the initial program for 5G-Advanced takes shape in 3GPP, this paper provides an overview of the key technology enhancements to be expected.

Contents

Introduction

The world's first 5G New Radio (NR) solution was standardized by 3GPP in Release 15. NR is highly flexible and enables significant improvements in three main service domains: enhanced mobile broadband (eMBB), ultra-reliable low latency communications (URLLC), and massive machine type communications (mMTC). These are achieved by a new User Equipment (UE), Radio Access Network (RAN) and 5G Core (5GC) design that includes numerous innovations. Among others, the RAN includes a new beam-based physical layer design with scalable numerology that allows operation up to mmWave carrier frequencies and supports massive MIMO, new channel coding schemes, and flexible frame structures. The Layer-2 RAN protocol design includes decoupling of higher layer functions from real-time constraints, avoiding duplicated functions, supporting front-haul interface splits, and flexible QoS. Among others, this is achieved by moving functionalities such as packet reordering and concatenation from the Radio Link Control (RLC) protocol layer to the Packet Data Convergence Protocol (PDCP), meaning that the tasks of the PDCP and RLC layers can be completed before fast scheduling decisions are made at the Medium Access Control (MAC) to trigger fast transmissions at the physical layer. 5G NR saw its first commercial deployments in 2019, using dual connectivity with LTE and the Evolved Packet Core (EPC) core network (so-called Non-Standalone (NSA) NR deployments), with the full 5G experience including the 5GC becoming available from 2020 when Standalone (SA) 5G networks started to be rolled out. 5G SA also enables deployment of private 5G networks for verticals.

Since the creation of 3GPP NR Release 15, the 5G standard has further evolved in Releases 16 and 17, introducing enhancements particularly for Industrial IoT (IIoT) and wider expansion of the 5G ecosystem, including innovations to support time sensitive communication (TSC), enhanced MIMO, small data transmission (SDT), UE energy saving, and many more. 5G evolution is now entering the 5G-Advanced era, starting with 3GPP Release 18. In July 2021, 3GPP held its first RAN workshop where potential innovations for 5G-Advanced were discussed, with companies submitting more than 500 proposals. In September 2021, 3GPP also discussed numerous potential innovations for the 5G-Advanced system architecture and services. In December-2021, 3GPP agreed on the scope of many Release 18 study and work items, while additional Release 18 items for UE and base station RF and performance requirements are expected to be decided in March-2022.

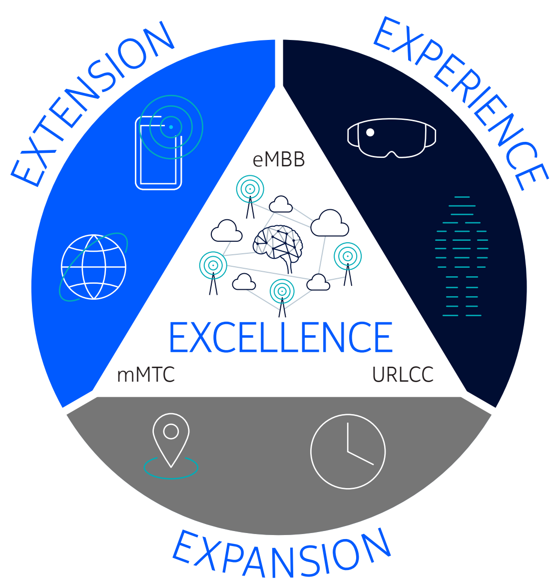
The first 5G-Advanced networks are expected to be deployed commercially around 2025. 5G-Advanced will provide a foundation for more demanding applications such as truly mobile extended reality services. It will also inject more intelligence into the network, utilizing machine learning (ML) to adapt to its environment. As with previous 5G releases, 5G-Advanced will boost fundamental radio and system performance, but it will also bring mobile broadband to new classes of devices and enhance support for novel use cases. As will be described in greater detail in the forthcoming sections, device innovations for 5G-Advanced are projected to include reduced capability devices, new sidelink innovations, reduced bandwidth operation, and eventually also making device satellite connectivity a commonplace feature. 5G-Advanced will be fully backward compatible, so it will be able to coexist with the current 5G NR Releases 15-17, including the ability to serve legacy 5G UEs.

Despite being still in the early phase of defining 5G-Advanced, in this paper we present insights into possible candidate innovations for 5G-Advanced. These will be motivated by the latest technology developments from academia and industry, as well as business drivers for enabling support for new use cases. As the number of innovations in 5G-Advanced is expected to be many, the description of each presented here is necessarily compact and is focused on explaining the motivations and likely benefits. As 3GPP provide little guidance on how all the 5G-Advanced items will combine to offer benefits for individual service provides, the paper is focused on presenting how 5G-Advanced is set to offer superior experience, expansions, extensions and operational excellence.

5G-Advanced narrative and timeline

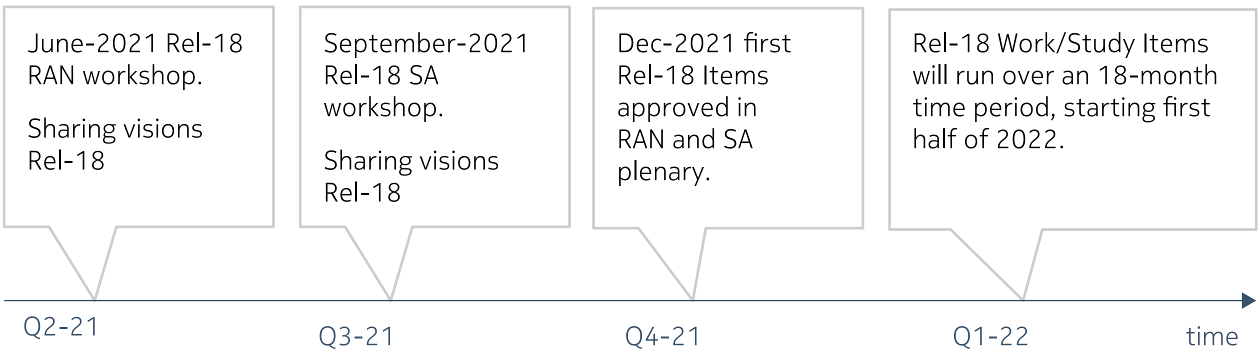
As pictured in Fig. 1, 5G-Advanced is expected to offer improved Experience for people and machines, Extensions for new use cases, and Expansions to offer new services beyond pure communication. This is powered by innovations that provide operational Excellence. The domain of enhanced Experience aims to lift 5G end-user experience to the next level, including better support for eXtended Reality (XR); enhancement techniques include further development of MIMO, and improvements in mobility and flexible duplexing are envisioned. The Extensions domain aims to extend the reach of 5G connectivity and to make it available to new market segments; including innovations for improved coverage, enhanced low-cost massive IoT, and further support for non-terrestrial networks (NTN) and drones. The Expansions domain targets the expansion of 5G services beyond traditional communication, by introducing enhanced positioning with sub-10cm accuracy consistently both indoors and outdoors, as well as time synchronization as a service, offering valuable benefits for use cases as diverse as smart power grid control, industrial automation and real-time financial transactions. 5G-Advanced will be powered by operational Excellence that aims to enhance and optimize the 5G platform and its operation by the gradual introduction of Artificial Intelligence (AI) and Machine Learning (ML) enablers, network slicing enhancements, wireline and wireless convergence, network coordination and energy efficiency enhancements. Energy efficiency improvements for both the network infrastructure and the devices will be in focus. These operational enhancements will ensure efficient network operation at affordable operational expense (OPEX) so that 5G-Advanced can efficiently serve a larger number of services with diverse QoS requirements.

Figure 1. Overview of the four innovation domains of 5G-Advanced.



5G-Advanced is expected to offer enhancements for both of the frequency ranges supported by 5G, namely Frequency Range 1 (FR1) that covers up to 7.125 GHz and Frequency Range 2 (FR2) that covers mmWave frequencies up to 71 GHz. The timeline for the first 5G-Advanced release, namely 3GPP Release 18, is shown in Fig. 2. The preparation for 5G-Advanced was ongoing during 2021, leading to specification work starting in 2022 and reaching completion around the end of 2023 and start of 2024 for the various items. Notice that although majority of the Release 18 items for RAN and SA were agreed at the December-2021 plenaries, additional Release 18 items for UE and base station RF and performance requirements are expected to be agreed in March 2022.

Figure 2. Timeline for 3GPP NR Rel-18: the first 5G-Advanced release.



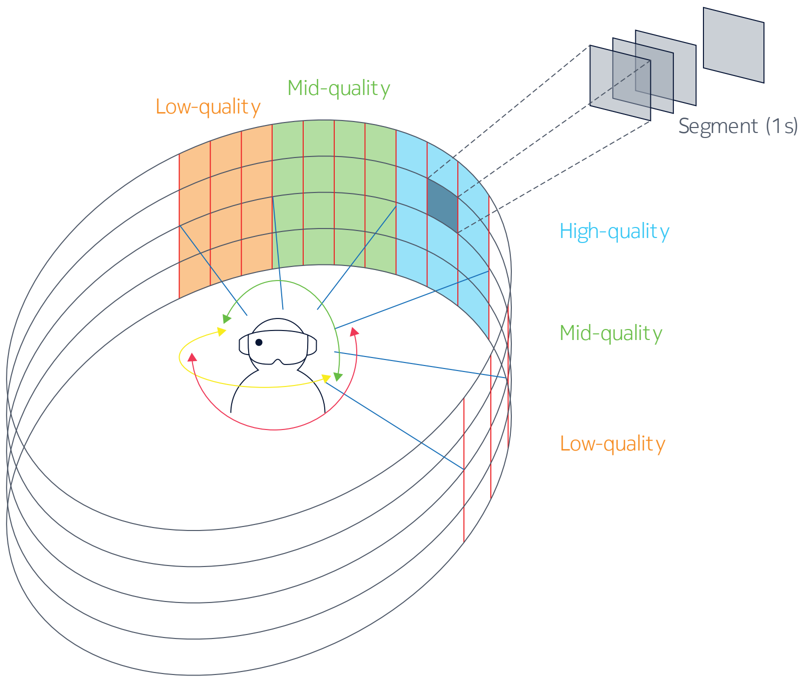
Experience

In this section we present the envisioned innovations that will lift the 5G end-user experience to the next level.

eXtended Reality and Media Services

XR is an emerging service built around combined real and virtual environments with associated human-machine interactions generated by computer technology and wearables. XR is an umbrella term covering augmented reality (AR), mixed reality (MR), virtual reality (VR) and cloud gaming (CG). It requires high data rates with strictly bounded latency constraints, typically in the order of 30-100 Mbps at the application layer, with packet delay budgets of 5-15 ms with a low packet error rate of 99% to 99.9%. The key to achieving a high level of support for XR in 5G-Advanced will be to provide the right quality of service for the different elements of XR data flows in a way that uses network capacity efficiently. One of the most important elements for XR applications is high-quality video transmission, often applying compression through a mixture of intra-frame and inter-frame coding. Mixed Reality (MR) uses pose information with up to six degrees of freedom (DoF) (position (x, y, z) and rotation (yaw, pitch, roll)) to minimize data rates by sending only the visible field of view; this is known as view-port dependent VR streaming and is illustrated in Fig. 3. Data rate can be adapted according to the accuracy of the pose information. In order to enable efficient Radio Resource Management (RRM) and scheduling of XR services, it is important that the traffic characteristics of the application are conveyed to the RAN. This calls for an improved 5G QoS framework, including extension of the 5G QoS Identifiers (5QI) that signal QoS characteristics from the CN to the RAN. As an example, six Degrees of Freedom (DoF) information may be included as a new application characteristic for QoS Flows and 5QIs defined for interactive XR and gaming services. This would enable view-port dependent bitrate adaption and scheduling optimizations, resulting in more efficient radio resource utilization and hence improving the network capacity in terms of the number of XR users that can be simultaneously supported with improved end-user experience.

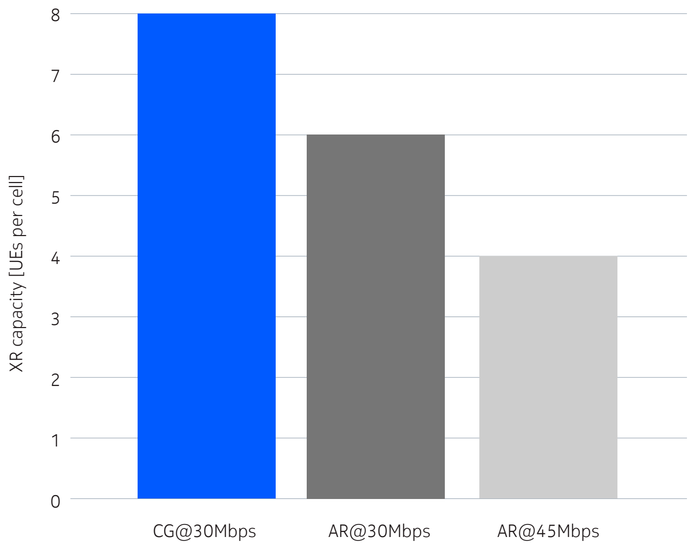
Figure 3. Example of view-port dependent virtual reality streaming.



The baseline performance for XR services and potential XR enhancements are quantified in a 3GPP study whose findings are captured in 3GPP TR 38.838. A frame-based XR traffic model is derived with a semi-deterministic arrival rate (typically 60 frames per second), including modest variations of the payload sizes. The time-domain jitter of frame arrivals and payload sizes are modeled with truncated Gaussian distributions.

Fig. 4 shows how the supported XR capacity in NR varies for three different cases. The XR capacity is defined as the maximum number of satisfied users, where a user is satisfied if 99% of the packets are correctly received within the packet delay budget (PDB); in Fig. 4, the PDB is 15 ms for CG and 10 ms for AR. The 30 Mbps case corresponds to a single stream of Full High Definition (FHD) video quality, while the 45 Mbps case uses 4K video. The deployment environment is dense urban, with a 100 MHz NR carrier in the 3.5 GHz band. As expected, the number of supported satisfied XR users decreases when comparing CG (with PDB of 15 ms) versus AR (with PDB of 10ms) with 30Mbps. As expected, if the AR data rate is increased to 45Mbps, the number of supported XR users is further reduced.

Figure 4. XR capacity for CG and AR in a Dense Urban scenario at 3.5 GHz with 100 MHz NR carrier bandwidth.



To further boost the XR capacity in 5G-Advanced, enhanced semi-persistent scheduling (SPS) has been identified as a promising feature as it offloads the otherwise busy dynamic MAC-based scheduler and helps save control channel overhead. Several SPS enhancements were already included in earlier 3GPP NR releases to serve Industrial IoT (IIoT) applications such as time-sensitive communication. However, it has been recognized that the current NR SPS does not support XR services well; for example, today's SPS cannot match the XR frame arrival rates, nor does it support efficient mechanisms to cope with variations of the frame size payloads. It is therefore expected that this will be improved for 5G-Advanced.

UE power consumption is also obviously of importance for XR services with a wide range of different types and form factors of XR devices, which may have very limited space for a battery and at the same time require demanding data rates. It is therefore expected that the Connected Mode Discontinuous Reception (CDRX) will be extended in 5G-Advanced to better match the XR traffic characteristics. The current estimate is that the device power consumption can be reduced by at least 20% for the XR use cases by introducing such enhancements, without jeopardizing the quality of experience.

Mobility and Cell Management

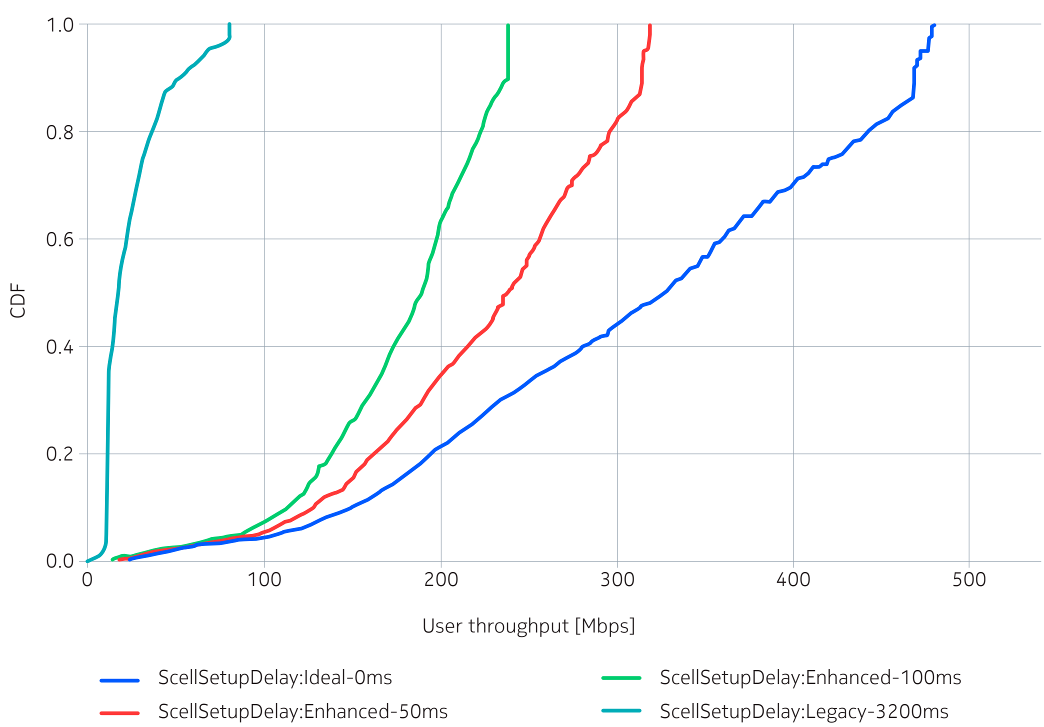
5G primarily relies on UE-assisted, network-controlled inter-cell mobility. It includes traditional handovers, where the network informs the UE of handover by means of RRC signaling, Conditional Handovers (CHO), where the UE initiates a handover on its own based on rules configured by the network, and Dual Active Protocol Stack (DAPS) handovers, which require the UE simultaneously to receive data from the source and target cells. These are all efficient handover mechanisms with low probability of handover failure, but DAPS comes at the cost of increased device complexity.

DAPS handover was introduced to reduce the interruption time for FR1-FR1, FR1-FR2, FR2-FR1 scenarios, but not for FR2-FR2 due to the device complexity implications. According to current NR requirements, the experienced handover interruption for DAPS is of the order of 1 ms for the downlink and 10 ms for the uplink when early forwarding of packets is assumed from the source cell to the target cell. With current DAPS handover, there is inherent end-to-end uplink latency of at least 10 ms while two signaling messages are sent over the Xn interface between base stations, before data forwarding to the User Plane Function (UPF) in the CN can start. Achieving close to 0 ms interruption time during handover in both downlink and uplink is desirable for 5G-Advanced as it would allow services to be deployed with stricter latency and reliability requirements (e.g. for IIoT and XR). Moreover, it would be desirable to be able to achieve this with only a single transceiver in the UEs, as some device types may have limitations in terms of size and power consumption.

Another aspect of mobility is efficient cell management for cases with Carrier Aggregation (CA) and/or Dual Connectivity (DC), to identify and configure additional secondary cell(s) in a timely way. The basic NR FR2 CA and DC cell identification, measurement and measurement reporting delays are significant; this degrades in user throughput in FR2 related CA and DC deployments, as additional high-capacity FR2 carriers cannot be configured quickly enough when data traffic is bursty. FR1 CA and DC performance is much better thanks to enhancements whereby measurements made while a UE was in Idle mode can be reported quickly upon entering Connected Mode without incurring further measurement delay, thus facilitating rapid secondary cell (SCell) activation. In FR2, the right beam configuration is critical, yet radio conditions in FR2 may change rapidly even at low-to-moderate UE speeds; FR2 Idle mode measurements reported in Connected mode may therefore not work as well for FR2 as for FR1. Improvements for FR2 SCell setup delays are therefore desirable.

The benefits of such early measurements and reporting enhancements in FR2 are illustrated in Fig. 5, showing the cumulative distribution function (cdf) for the downlink end-user experienced throughput. These results were obtained using 3GPP-compliant dynamic simulations with time-varying FTP Model-3 traffic, assuming a deployment with a 20 MHz FR1 carrier and a 100 MHz FR2 carrier. From these results, it is clear that the basic NR FR2 SCell setup latencies cause significant degradation in experienced user throughput in CA and DC deployments. The performance is clearly improved if enhanced early measurement and reporting procedures are introduced for FR2 to enable faster SCell activation.

Figure 5. Cumulative distribution function (cdf) of end-user experienced data rate under different assumptions for the delays of FR2 SCell setup.



Tighter integration of intra-cell beam management (including UEs with multiple antenna panels) is an area of possible improvement for 5G-Advanced that is especially relevant for FR2. As an example, the current UE beam refinement procedure is performed only after handover execution is completed. Performing UE beam refinement at this late stage causes additional signaling overhead and increases the delay to set up a narrow beam for the UE by 15-20 ms, and this could be improved by integrating the UE beam refinement into the handover procedure. Similarly, enabling UE beam refinement for a CHO candidate target cell before the CHO execution is performed is also desirable. All of this would essentially make handovers transparent to the experienced QoS even for demanding applications with high data rates and tight latency and reliability constraints.

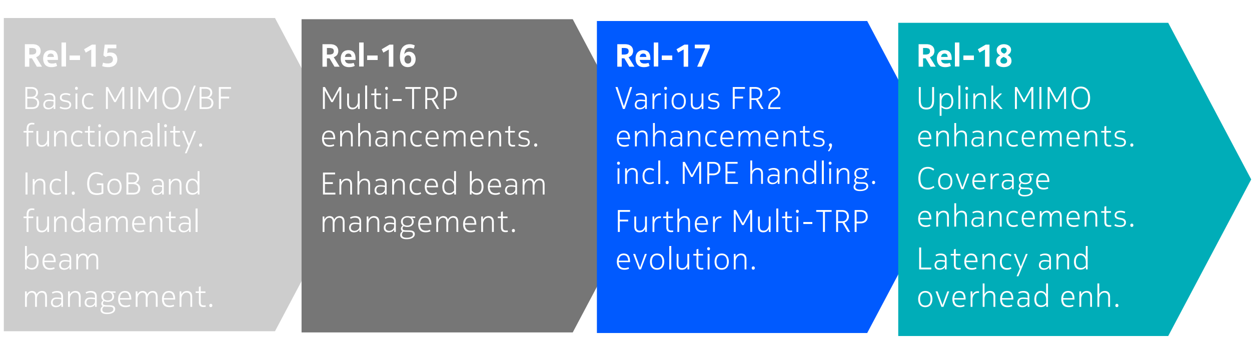
Additionally, Maximum Permitted Exposure (MPE) events at the UE can result in the transmit power having to be reduced, and such events can impact mobility performance if not taken properly into account. Awareness of the UE MPE status at both the source and the target cells would therefore be advantageous, to reduce handover failures that could otherwise be caused by MPE limitations.

Uplink coverage and MIMO/beamforming enhancements

The physical and MAC layers of 5G were designed for advanced MIMO and beamforming at both the base stations (gNBs) and the UEs from the start, including suitable reference signals, channel state information (CSI) feedback and MAC-based beam management procedures. The evolution of MIMO and beamforming techniques will continue in 5G-Advanced. While the downlink MIMO schemes have been extensively addressed in the first releases of NR, 5G-Advanced is expected to put more emphasis on uplink enhancements, with higher rank and multilayer transmissions, including support for UEs with more uplink transmission chains. This will be beneficial for standalone operation where UE transmission resources are fully available for 5G usage thus enabling uplink MIMO. We predict that such techniques could bring 20% uplink throughput gain for 5G-Advanced as compared to 5G. Fundamental uplink coverage enhancements are also expected, where the peak-to-average power ratio of higher order modulation (QPSK and 16QAM) transmissions could be reduced by up to 2 dB by applying frequency domain spectrum shaping (FDSS) with spectrum extension. This would enable higher data rates at the cell-edge.

More sophisticated Multi-TRP (transmission and reception point) enhancements are envisioned, such as multi-codeword encoding for Multi-TRP transmission with a single downlink control information (DCI) message. Also, multi-beam enhancements such as enablers for dynamic configuration of separate transmission configuration indicators (TCIs) and latency and overhead reductions for UE beam management are expected. These will lead to improvements in spectral efficiency and robustness of signaling. Fig. 6 shows a summary of the main MIMO/beamforming features from 5G legacy Release 15 to 5G-Advanced Release 18.

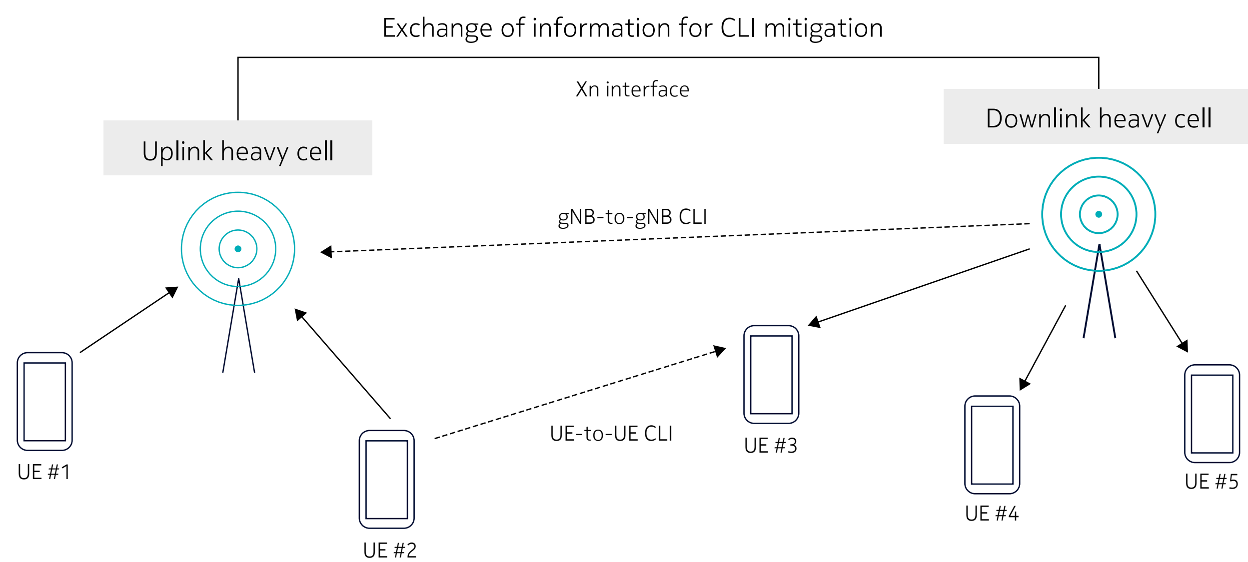
Figure 6. Summary of the main MIMO/beamforming features in different releases.



Flexible Duplexing

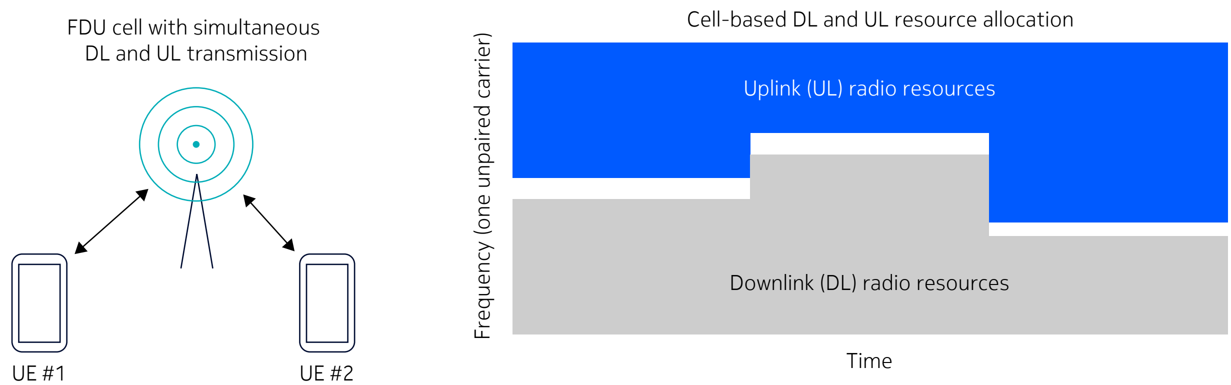
The flexible frame structure design of NR supports Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD) for paired and unpaired carrier deployments respectively. Some enhancements for TDD operation were introduced in Release 16, including: (i) remote interference mitigation (RIM) between cells separated by up to hundreds of kilometers, and (ii) a framework for simple coordination of TDD radio frame configurations between network elements, with UEs measuring UE-to-UE cross link interference (CLI). However, standardized solutions for management of gNB-to-gNB CLI are currently missing, and hence on the table for 5G-Advanced. Such solutions are of particular importance for enabling so-called uplink-heavy configurations, which are becoming increasingly important for network operators as the ratio between uplink and downlink traffic shifts. Fig. 7 illustrates the gNB-to-gNB and UE-to-UE CLI which arises in a scenario in which a (victim) cell is using a TDD configuration mainly allocated to uplink traffic, while being subject to interference from an (aggressor) cell that operates with a downlink-heavy TDD configuration. Options for mitigation of gNB-to-gNB CLI include: (i) coordinated time-frequency muting, (ii) coordinated beamforming, (iii) advanced gNB receivers with enhanced gNB-to-gNB CLI suppression or cancellation capabilities. As an example, the downlink-heavy cell may offer assistance information to the uplink-heavy cell to enhance its gNB-to-gNB CLI mitigation performance.

Figure 7. Interference paths for different flexible duplexing options.



Potential future adoption of Full Duplex (FD) solutions in 5G-Advanced is also under discussion. FD basically enables simultaneous uplink and downlink transmission on the same radio resources, relying on advanced gNB and UE self-interference cancellation schemes to cope with strong downlink-uplink and uplink-downlink interference at the network nodes and devices. In theory, FD has the potential to offer improvements in latency, coverage, and capacity compared to traditional TDD. The degree of achievable gNB and UE self-interference cancellation, as well as the complexity, energy consumption and cost, are expected to be determining factors for potential future integration of FD in 5G-Advanced. However, as an intermediate step, solutions where simultaneous uplink and downlink transmission occurs on non-overlapping radio resources within one unpaired carrier will first be further developed as is pictured in Fig. 8. This is known as flexible duplexing (FDU), where the division between downlink and uplink assigned radio resources may be dynamically adjusted per cell. The hypothesis is that such solutions will offer similar advantages as FD by enabling simultaneous uplink and downlink transmissions, but at a potential lower cost of self-interference cancellation (as uplink and downlink transmissions are non-overlapping in the frequency domain).

Figure 8. Simple illustration of a cell with FDU, where the gNB has simultaneous DL and UL transmission and reception on the same unpaired carrier on non-overlapping radio resources.



Finally, inter-operator coexistence aspects need to be carefully considered, as well as coexistence of 5G NR legacy UEs without advanced flexible duplexing capabilities. Notice that the Release-16 TDD coexistence studies concluded that dynamic TDD is primarily applicable for small cell nodes, while high power macro sites should use fixed TDD configurations to avoid performance degradation.

Edge Computing enhancements

As the speed of light is limited to 300,000 km/s (and only 200,000 km/s in fiber), ultra-low latency on the user plane requires that the two endpoints of the communication path, namely the application client on the device side and Application Server (AS) on the network side, are close to each other. Ultra-low latency applications therefore cannot be hosted in the internet with central peering point, like most of today's applications, but need to be hosted in so-called Edge Clouds (sometimes referred to as Edge Computing). Edge computing allows operators to host applications and content closer to the user, thus ensuring low latency for edge applications while keeping heavy traffic at the network edge and away from the backbone network.

Edge computing is supported as an integral part of the 5G System architecture. The 5G system includes functionalities such as an uplink classifier (UL CL) and multi-homing to enable UEs to obtain access to content available locally in the edge application server while also retaining the ability to keep the IP address and service anchored in a central User Plane function to enable seamless service continuity. 5G also introduces a new mode of service continuity with make-before-break relocation of the UPF for mobile UEs and enablers for application functions to influence traffic routing (e.g. the application function (AF) can indicate which traffic is to be offloaded, and for which UE). The 5G system also supports functionalities such as dynamic insertion of UL CL based on user plane traffic from the application, and the ability for the operator to use UE route selection policies (URSPs) to map certain application traffic to certain Protocol Data Unit (PDU) sessions. A PDU Session is an association between the UE and the Data Network, it can be of type IPv4, IPv6, IPv4v6, Ethernet, or Unstructured.

For 5G-Advanced, the following edge computing enhancements are considered:

- Exposure of radio conditions averaged over a period of time towards the application for a given UE.
- Enabling different policies for different categories of UE (without having to create pre-defined static groups of numerous UEs) and the ability for the AF to influence the same based on application usage.
- Ability to relocate a given edge application server for a collection of UE(s) while it is not required for them to be members of a pre-defined group (e.g. for multi-user gaming).
- Potential enhancements for edge computing in roaming situations; traffic offload for home-routed PDU sessions is currently not supported as it is especially tricky because home-routed PDU Sessions are meant to keep the Home Control in full control of the traffic (e.g. for lawful intercept and charging purposes).

The main benefits of these edge computing enhancements for 5G System are to enable efficient delivery of certain services from operators or 3rd party content providers, as well as to provision the services over long distances in a cost-effective way. This enables access to services deployed close to the UE's point of attachment, while at the same time offering access to regular internet, in a flexible and cost-efficient manner that can also leverage Network Function Virtualization (NFV).

Extension

In this section we present the innovations that aim to extend the reach of 5G connectivity, as well as to extend it to new market segments. The focus is on use cases which require connectivity for IoT and mobile broadband, including extended battery life, extended coverage, and low cost.

Reduced Capability devices

As the 5G ecosystem expands, new types of device are emerging that can benefit from the coverage and efficient connectivity of 5G, yet do not need to implement every aspect of its immense performance and flexibility. Such devices were first standardized in Release 17, known as Reduced Capability (RedCap) NR devices (initially also known as NR-light). These are designed to be much cheaper than full-NR devices, with, for example, fewer receive antennas and reduced RF bandwidth (only 20 MHz instead of 100 MHz).

It is now evident that further reductions in complexity would open new ecosystem domains to the benefits of 5G networks, especially in low-end video surveillance devices for industrial quality control and process monitoring, as well as in sensing and asset tracking. Such devices would benefit from improved power efficiency, and 5G-Advanced is therefore expected to introduce lower power devices that can be designed with on-chip power amplifiers, enhanced sleep modes, and possibly even the ability to be powered by energy harvesting, for which devices have to be able to operate from energy that is available only in small amounts for short periods. Devices with further reduced baseband processing and potentially RF bandwidth; 5 MHz seems to be a sweet-spot in terms of minimizing device cost, offering suitable data rates for the targeted applications, and at the same time avoiding redesign of fundamental aspects of the 5G radio interface, such as the synchronization and broadcast signaling. Fig. 9 shows a prediction of how RedCap devices in Releases 17 and 18 (R17 and R18) compare to legacy 5G mobile broadband (MBB) devices in terms of bandwidth capabilities, peak data rates and costs.

Figure 9. Comparison of 5G-Advanced RedCap device characteristics versus legacy 5G MBB devices.

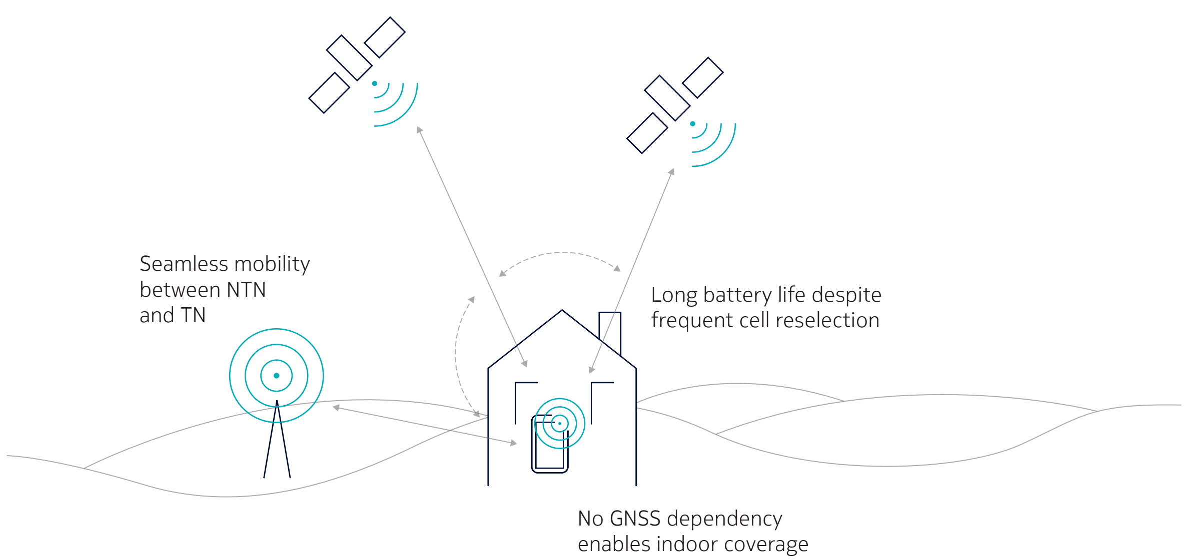
	5G MBB	Red Cap R17	Red Cap R18
Bandwidth	100 MHz	20 MHz	5 MHz
Peak rate	2 Gbps	100 Mbps	10 Mbps
RF cost	Reference	-44%	-46%
Baseband cost	Reference	-70%	-89%
Total cost	Reference	-60%	-71%

RedCap devices, as well as higher-capability 5G devices, may often engage in short transmissions of small amounts of data, and these need to be handled efficiently. 5G includes specific features for such "Small Data Transmission" (SDT) communications, and further optimizations of these procedures are envisaged for 5G-Advanced. These include enhanced support for mobile-terminated SDT, backhaul enhancements to avoid unnecessary fetching of the UE context, and improvements in reliability.

Non-terrestrial communications

5G NR Release 17 for the first time supports non-terrestrial networks (NTN) using satellites and high-altitude platform stations (HAPS). A key benefit of such networks is enhancement of coverage for truly global service provision. Full integration of NTNs with the conventional terrestrial NR networks is expected to come in 5G-Advanced. This will enable full mobility and service continuity as UEs roam from areas with good coverage from terrestrial base stations to more remote areas. Early implementations of satellite-based NR NTNs require devices to be synchronized to a Global Navigation Satellite System (GNSS) system in order to facilitate estimation of timing advance and pre-compensation of the high Doppler shifts inherent in low-earth-orbit (LEO) satellite communications. Further evolution of 5G-Advanced with potential elimination of this dependency would enable indoor connectivity for NR NTNs. Improvements in link budget are also expected, leading in due course to the possibility of NR NTN connectivity becoming a commonplace smartphone feature. These developments are illustrated in Fig. 10.

Figure 10. Non-terrestrial network use cases and characteristics for 5G-Advanced.



Closer to the ground, Unmanned Aerial Vehicles (UAVs), or drones, are expected to proliferate during the timeframe of 5G-Advanced, with applications ranging from search and rescue, to parcel delivery, to aerial displays. While 5G connectivity for UAVs has been possible since the beginning of NR, dedicated enhancements would significantly improve the efficiency of UAV communication and control. Specifically, enhancements in measurement reporting, triggered by the altitude of the UAV, and indications of current and future flight paths, would improve the ability of the NR network to manage these communications. Subscription-based identification of aerial UEs will also be important for security purposes. Also possible broadcast of UE identification will be considered.

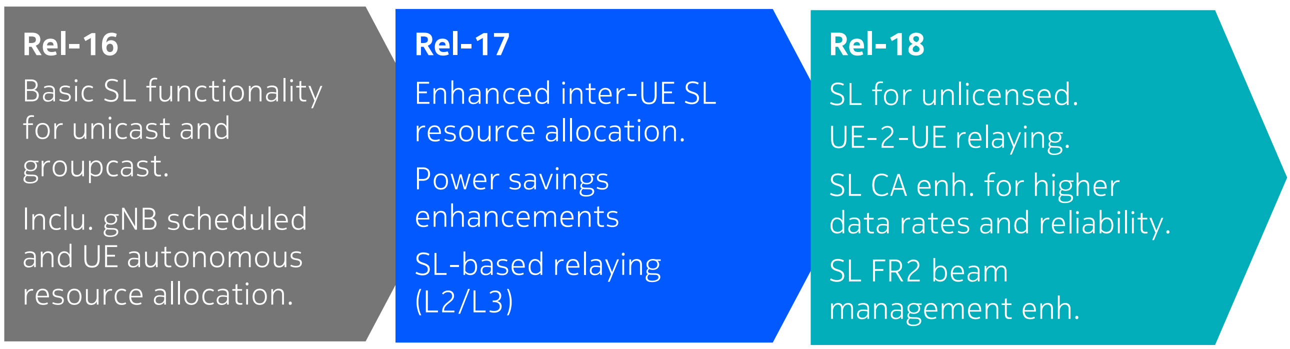
Dedicated radio beam management techniques that are likely to be employed in 5G-Advanced can be very beneficial with UAV communications, since a UAV may have strong line-of-sight to different base stations from those on the ground in the direct vicinity of the UAV. This can be both a curse and a blessing. On the one hand, a UAV may cause interference to relatively distant base stations, and this may need to be mitigated by beam control. On the other hand, in a scenario such as a UAV-mounted TV camera surveying a sporting event in a stadium, the UAV may be able to direct a beam to a base station outside the stadium that is not visible to the crowds on the ground; this could provide a reliable communication link for the high-definition video stream from the UAV, while leaving the capacity of the base stations in the stadium itself for the benefit of the spectators.

Sidelink enhancements

Direct UE-UE communication via the sidelink was introduced into NR in Release 16, focusing first on the substantial market for automotive communications, often referred to as V2X (Vehicle-to-vehicle/pedestrian/network, etc.). This supports both basic automotive safety use cases and more advanced applications such as automated driving. Sidelink communication is also important for public safety services, such as emergency responders being able to communicate with each other beyond network coverage. Release 17 includes some enhancements for the latter case, including UE power-saving measures and UE-network relays to extend network coverage in emergency zones, as well as resource scheduling enhancements to improve reliability for V2X use cases.

5G-Advanced is expected to build on the sidelink foundation of NR from earlier releases in several directions. Firstly, sidelink operation will be extended to unlicensed spectrum with corresponding channel access procedures. This will address the currently limited spectrum availability for sidelink operation. Secondly, the relaying functionality will be extended to include UE-UE relays, which open the way to multi-hop relaying, for example to bring coverage deep into disaster zones. Thirdly, higher sidelink data rates may be supported by means of carrier aggregation, enabling more advanced V2X applications such as “see through” video to operate with low latency, giving drivers greater awareness of road conditions further ahead. Sidelink carrier aggregation will also enable improved reliability by using packet duplication on the different component carriers. Fourthly, sidelink enhancements for FR2 to improve the beam management performance are expected. Finally, sidelink use cases are foreseen related IIoT applications, and extensions to those are likely to be developed. A summary of the main sidelink features from legacy 5G Release 16 to 5G-Advanced Release 18 is pictured in Fig. 11.

Figure 11. Overview of the main sidelink features in different releases.



NR support for dedicated spectrum less than 5 MHz

Extending 5G services to certain specialized vertical use cases calls for the ability to deploy NR in dedicated spectrum bandwidths below 5MHz. For example, in Europe it has been decided that the Future Railway Mobile Communication System (FRMCS) which will replace GSM-R will use NR. Soft migration from GSM-R to FRMCS is expected to take place between around 2025 and 2035; this will necessitate co-existence of GSM-R and NR within 2x5.6 MHz of FDD spectrum, leaving approximately 3.6 MHz for NR operation. Other similar use cases include smart grids (2x3 MHz FDD in US) and public safety (2x3 MHz FDD in Europe).

Unlike the RedCap use cases, these vertical use cases do not impose constraints on device complexity. The goal, therefore, is to adapt NR to the available dedicated spectrum allocations with minimal changes, building on the existing ecosystem of NR devices and infrastructure. To this end, the synchronization signals should not be changed (since their bandwidth is less than 2 MHz), while straightforward puncturing of the edges of channels such as the Physical Broadcast Channel (PBCH) and Physical Downlink Control Channel (PDCCH) is likely to be the preferred course of action. In terms of RF requirements, existing filtering assumptions from the 5 MHz NR channel bandwidth are likely to be retained for the range of transmitted bandwidths required for FRMCS, since the same operator would be deploying the adjacent GSM-R and FRMCS systems, thus allowing the co-existence aspects to be managed. A new channel bandwidth of around 3 MHz is likely to be specified for the smaller dedicated spectrum bandwidths for the other use cases.

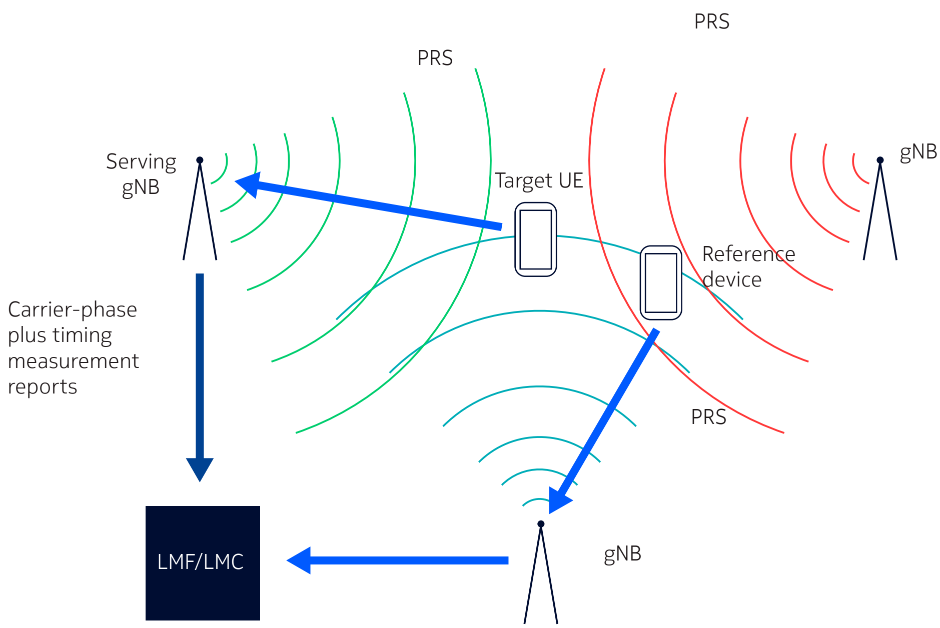
Expansion

Innovations that offer expansions beyond traditional communication services are presented next.

Positioning enhancements

Cellular-based positioning was introduced in 5G NR in Release 16. The wide bandwidths and beamforming capabilities of NR enable it to support a range of positioning techniques that can be combined to give an accurate result: downlink Time Difference of Arrival (TDOA), uplink TDOA, downlink Angle of Departure (AoD), uplink Angle of Arrival (AoA) and multi-cell Round trip time (RTT), as well as enhanced cell ID. Enhancements in Release 17 include improved identification of non-line-of-sight paths, which are one of the main sources of uncertainty in positioning estimation. The 5G network can also aid signaling for real-time kinematic (RTK) GNSS based on GNSS carrier phase estimation, which can provide centimeter-level accuracy, but only outdoors.

Figure 12. Carrier phase positioning applied to terrestrial base stations.



5G-Advanced is expected to deliver a further step-change in positioning accuracy, especially indoors where GNSS is unavailable. Much research on potential enhancement techniques has been undertaken in recent years. The most promising technique for such accuracy enhancements is carrier-phase positioning applied to reference signals transmitted by the 5G-Advanced base stations, as shown in Fig. 12. The basic technique is already well known from RTK GNSS positioning, and it should enable accuracies of a few centimeters to be achieved consistently both indoors and outdoors.

Sidelink-based ranging is also expected to be included in 5G-Advanced, either as a standalone technique for relative positioning or in conjunction with other techniques to enhance absolute positioning. Further enhancements to positioning in 5G-Advanced are likely to focus on reducing the power consumption and latency in obtaining a positioning estimate, for example by enabling positioning while in Idle mode.

Timing Resiliency and Synchronization as a Service

Telecom networks usually have a combination of synchronization methods (e.g., local GNSS receiver at RAN nodes, Precision Time Protocol (PTP) transport network, physical layer clock) to ensure reliability and robustness of phase/time synchronization (i.e. to avoid RAN timing failure). GNSS receivers determine the precise time by processing the signals broadcast by satellites. PTP is a protocol that is used to synchronize time across networks.

With the evolution of 5G networks, more stringent requirements on network synchronization lead to gNBs having the timing resiliency capabilities (e.g., high frequency stability, holdover capabilities) that are needed for when the primary source of synchronization is unavailable. Hence, 5G-Advanced will be in a unique position to provide a general timing resiliency system that can coexist with the widely used timing standards based on IEEE, IEC, and ITU-T protocols, and it will be able to supplement vendor-specific solutions.

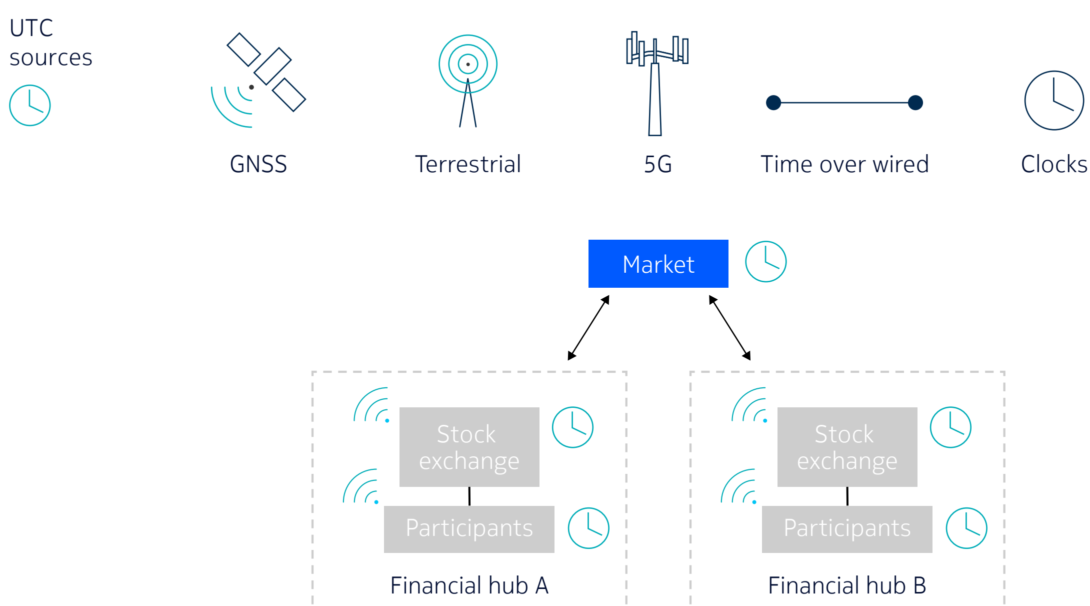
The 5G system can act as an intelligent timing resiliency management system that can ease the supervision of different time sources or equipment combined in the 5G deployment. The 5G core can collect time source status reports (e.g. GNSS reports from RAN nodes) and 5G system timing resiliency capabilities (e.g., holdover and frequency stability capabilities) to determine the status of the system in a wide area network and influence consistent operation of timing resiliency (e.g., amongst RAN nodes) in a certain service area. 5G timing resiliency management in 5G-Advanced will bring two main benefits:

i) 5G core can provide a consistent fault-tolerant mitigation solution for the area of interest (e.g., a whole operator's 5GS deployment, per service area, per PLMN).

ii) Timing resiliency capabilities awareness (e.g., clock diversity, holdover) can relax time synchronization dissemination requirements (e.g., to 5G network elements or end users). That is, time synchronization as a service should consider multiple dimensions to fine-tune the expected performance, such as the area of interest, accuracy, clock diversity and holdover requirements.

An example of applying 5G for timing resiliency purposes within the stock exchange for precise timestamping of financial events is shown in Fig. 13. Since 5G coverage is expected to be widely available, it can also be leveraged as an alternative source to offer timing as a service to other industries that are currently reliant on GNSS timing solutions.

Figure 13. Overview of time Synchronization as a Service (SaaS).



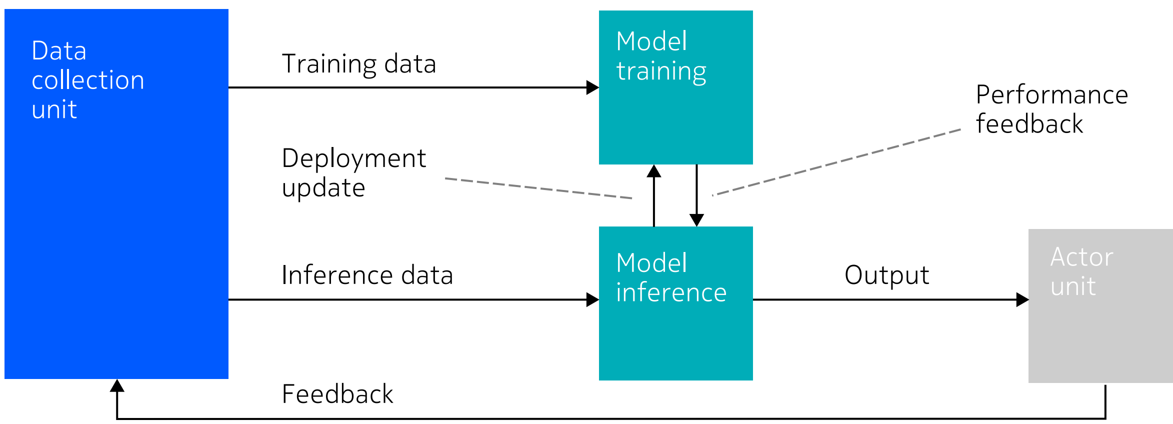
Operational excellence

In this section we present an overview of the envisioned innovations that aim to enhance and optimize the 5G platform and its operation in various ways.

AI/ML enablers

5G-Advanced will introduce new artificial intelligence (AI) and machine learning (ML) technologies across the RAN, core and management network domains. These technologies will help unlock the full potential of 5G systems, empowering them with new network automation capabilities, boosted performance and enhanced energy efficiency. AI/ML techniques are already applicable for today's 5G systems, primarily as implementation-specific innovations without explicit 3GPP standardization support. Examples include ML-driven RRM. 3GPP NR Release 17 include a study that aims at enabling RAN intelligence by defining the basic framework for AI/ML in the RAN in an agnostic way without modifying the RAN architecture and interfaces. This is considered for the purposes of energy saving, load balancing and mobility enhancements/traffic steering, among others. The Release 17 study is expected to recommend a functional framework to enable AI/ML in the RAN, while not yet covering UE impacts. Hence, for 5G-Advanced, there will most likely be standardization of harmonized data collection, offline model training and management, and inference functions such that AI/ML can be supported. Specifics of AI/ML algorithms will not be explicitly specified by 3GPP, but are expected to remain vendor-specific. The 3GPP generic functional framework for AI/ML RAN intelligence as captured in 3GPP TR 37.817 is pictured in Fig. 14. This will complement the network data analytics function (NWDAF) in the 5G core network, where AI/ML enablers were introduced in earlier releases.

Figure 14. 3GPP proposed generic framework for RAN intelligence.



Use of online AI/ML for the lower RAN layers such as PHY and MAC is also under discussion for 5G-Advanced studies. The open literature includes many such studies, where the use of various realizations of Reinforcement Learning (RL) solutions appear to be popular, especially for gNB online learning features (e.g. for RRM). 3GPP will undertake a study on AI/ML for the lower RAN layers for 5G-Advanced, starting with defining an updated link- and system-level simulation methodology that enables realistic performance studies of such AI/ML based techniques. Once a common methodology framework is established, candidate enhancements for AI/ML can be studied and benchmarked to clarify what potentially may be standardized to unleash the potential of such techniques for cases where real-world attractive gains are found. Initial use cases to be addressed for potentially applying AI/ML in 5G-Advanced include CSI feedback enhancements to achieve further overhead compression, potentially with elements of channel state prediction, predictive beam management and mobility, and positioning accuracy enhancements. Robustness, convergence properties, and complexity considerations will need to be considered as well, including how sufficient training (exploration) can be guaranteed.

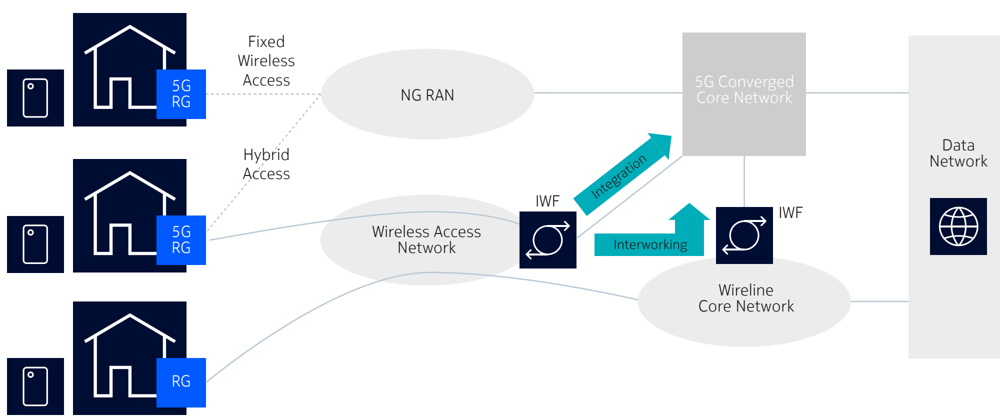
Finally, testability and minimum performance requirements for devices that may include AI/ML RAN functionalities must be studied and addressed, to ensure that the use of AI/ML in the devices brings benefits relative to the 5G benchmark.

Wireline and wireless convergence

The 5G System is designed to support multiple access technologies natively, including both 3GPP and non-3GPP access like Wi-Fi and wireline as illustrated in Fig. 15. The 5G Core can be described as access agnostic, as it offers seamless user experience when a device moves from one access technology to another. It enables common authentication, registration, mobility management and session management, and provides a universal policy framework for all access technologies. Using a common 5G Core for different access technologies makes it possible to maximize data rates and increase reliability, and thus to improve user experience: 5G supports traffic steering capabilities between wireless and wireline access, allowing selection of the best network for a given service data flow. Wireline access is considered as a trusted access (sometimes also referred to as trusted non-3GPP access) from the 5G System point of view, as wireline-wireless converged access is expected to be managed by the same operator.

5G System capabilities will be improved in the future to better serve devices connected via a 5G Residential Gateway (5G-RG), for example to provide differentiated IP services to different devices connected via a 5G-RG and to improve support of community WiFi services. Furthermore, 5G-Advanced will bring the ability to consider different subscription data when the 5G-RG is connecting to the network from different locations and to improve mobility support for UE(s) moving between trusted non-3GPP access networks.

Figure 15. Wireline and wireless convergence.



CU resiliency

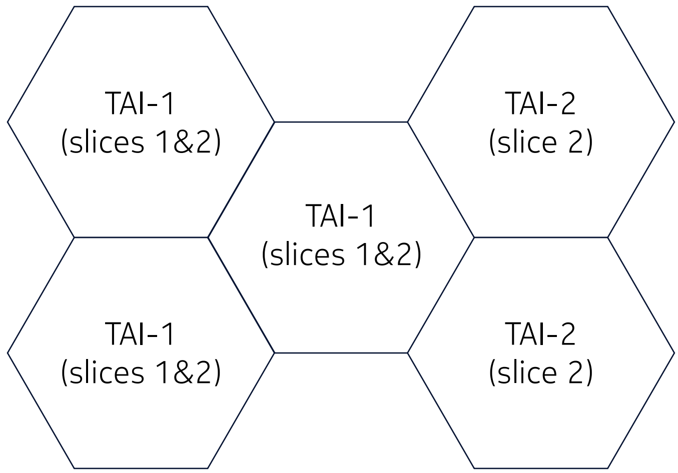
The network architecture for 5G is rather flexible. One of the options is the so-called split-gNB architecture that consists of a Centralized Unit (CU) that may serve up to 512 Distributed Units (DU). This offers opportunities for centralized RAN solutions for large clusters of 5G NR cells. As the CU hosts the RRC protocol, it essentially holds the responsibility for the control plane for all connections. 5G-Advanced is therefore expected to address additional resiliency mechanisms against CU failures. This will include evaluating the problems (e.g., failure detection, switchover) and performance impacts associated with such CU failures, and investigating potential solutions (e.g., cold stand-by, context transfer) to minimize downtime, user dropping, and performance impact in case of CU control plane failures. Geo-redundant resiliency solutions should also be investigated, while remaining compatible with the existing NG-RAN architecture and cardinality rules.

Network slicing

Network slicing is an end-to-end concept that enables isolation and optimization of part of the network for one or more service types or enterprises. A network slice spans the RAN, core network and transport network domains. It is a logical instantiation of a network that enables proper isolation with a rapid life cycle – 'fast to deploy', 'fast to decommission'. Network slicing is a powerful functionality that helps to manage highly diverse services over one 5G network and it is considered a mandatory functionality for the UE and network. Foundational enablers for network slicing, as well as specific improvements to support various business cases, are already present in the current 5G System (Release 15 to Release 17). Some of the enablers include network slice selection, network slice specific authentication, tiered service support (i.e., the ability to control the number of registered devices, number of sessions and data rate per slice for a given slice customer), band-specific network slices, slice-aware cell reselection and initial access as well as slice continuity during mobility events. Slice availability is provided in terms of tracking areas and UE is associated with certain slices are tracked by means of tracking areas (TAs) and registration areas (RAs). Until Release 17, it has been assumed that a certain slice is supported homogenously within the registration area which is comprised of one of more TA(s). Each tracking area is identified by a Tracking Area Identifier (TAI), which is broadcast by each cell. Innovations for ease of network operation and optimizations needed to support advanced business cases are expected to be added in 5G-Advanced, including:

- Minimization of service interruption when a slice service area border is crossed. This could be achieved by further extending the so-called slice re-mapping work of Release 17. More specifically, for scenarios where an existing slice cannot serve the application in the current cell (for OAM reasons) or the target cell (due to mobility), or if the existing slice cannot meet the performance requirements of the applications, system optimizations may be considered.
- Addressing deployment scenarios where network slice support is not uniform within a certain registration area. For example, a network slice may be supported in a certain TA but not in another TA within the same registration area, as shown Fig. 16 in which TAI-1 supports slices 1&2, while TAI-2 supports only slice 2 and not slice 1. 5G-Advanced is expected to provide a solution for the UE to initiate registration for a network slice even though it was rejected in one tracking area but available in another TA within the same registration area. This is to facilitate deployments that consider band specific slicing and mapping of certain services to certain bands, or to facilitate orchestrating a certain slice within a smaller area for a shorter period of time.
- Ability for the network to control when the UE registers, deregisters with certain network slices.
- Ability for a roaming UE to perform PLMN selection based on supported slices.

Figure 16. Non-homogenous support of Network slice.



Network energy efficiency

5G-Advanced will have a pivotal role in the global efforts towards sustainable development goals. Findings from early 5G deployments have shown that 5G is already up to 90% more energy efficient than legacy technologies, and significant improvements can be achieved with more energy-optimized implementations. However, as the exponential growth of cellular traffic continues, it is crucial that that energy efficiency continues to be addressed in 5G-Advanced. The challenge from the standardization perspective is twofold: (i) 5G-Advanced features should not compromise network energy efficiency, (ii) the 5G-Advanced standards should enable implementation of new energy saving features. As a first step to tackle these challenges, a common methodology to estimate the relative energy consumption of future networks needs to be developed. Existing energy consumption models introduced by 3GPP can be taken as a starting point for such studies. To tackle the second challenge, the potential of energy saving techniques will need to be studied in the spatial, time, frequency and power domains, as well as considering what role data collection and UE assistance may be able to play in network energy saving.

In order to assist the world with environmentally sustainable connectivity and low operating costs, 5G-Advanced will need a combination of well-designed standards and outstanding implementation. Those may likely build on methods that aim to utilize statistical load variations to obtain energy savings in line with the traffic demands. Also the use of AI/ML based techniques are expected to play an important role, where the framework for RAN intelligence presented earlier as pictured in Fig. 14 can be utilized to, for example, predict traffic variations to obtain further energy savings.

Concluding remarks and outlook

In this paper we have a provided a first outlook for possible key innovations that are likely to be introduced in 5G-Advanced, as summarized in Fig. 17. 5G-Advanced is set to evolve 5G System to its fullest capabilities, from 3GPP Release 18 onwards. The innovations from the large number of 3GPP 5G-Advanced items are foreseen to offer improved Experience for people and machines, Extensions for new services, and Expansions to offer new functionalities. It will be powered by technological innovations that provide operational Excellence. Among others, it will continue to improve coverage and capacity, enhance end user experience and expand 5G capabilities beyond connectivity. 5G-Advanced is also expected to inject more intelligence into the network, utilizing artificial intelligence and machine learning to adapt automatically to the environment, boost fundamental performance, bring mobile broadband to new classes of devices and open applicability to new industries. This will be achieved by introducing a substantial set of innovative enhancements on top of the existing 5G design, while maintaining backward compatibility to 5G. The innovative features outlined in this paper are all candidates for 5G-Advanced, based on the latest proposals and analysis from industry and academia. The 3GPP Release 18 specifications forming the start of 5G-Advanced are expected to be available early in 2024. While the scope of 3GPP Release 18 is starting to take shape, the research community continues to innovate for 5G-Advanced Releases 19 and 20 that eventually will provide a bridge into the future 6G era.

Figure 17. Summary of 5G-Advanced domains and related candidate features.

Extension	Extended reach and new segments	• Uplink coverage • Sidelink enhancements	• RedCap Evolution • Non-terrestrial networks • Drone optimization
Experience	Improved 5G experience	• Extended reality (XR) • Cloud gaming • Mobility performance	• Beamforming boost • Edge computing • Flexible duplexing
Expansion	Beyond traditional communication	• Timing resiliency • Synchronization service	• Accurate Positioning
Operational excellence	Improved operability	• Centralized unit resiliency • Data analytics with ML/AI	• Slicing enhancements • Network energy saving • Wireline wireless convergence

Abbreviations

3GPP	Third Generation Partnership Project
5G	Fifth Generation
5G-RG	5G Residential Gateway
5GS	5G System
5QI	5G QoS Identifier
AF	Application Function
AI	Artificial Intelligence
AoA	Angle-of-Arrival
AR	Augmented Reality
CA	Carrier Aggregation
CG	Cloud Gaming
CLI	Cross Link Interference
CHO	Conditional HandOver
CN	Core Network
CP	Cyclic Prefix
CSI	Channel State Information
CU	Central Unit
DAPS	Dual Active Protocol Stack
DC	Dual Connectivity
DCI	Downlink Control Information
DoF	Degrees of Freedom
DU	Distributed Unit
eMBB	Enhanced Mobile Broad Band
eMTC	Enhanced Machine Type Communications
FD	Full Duplex
FDD	Frequency Division Duplexing
FDU	Flexible DUplexing
FHD	Full High Definition
FR	Frequency Range
FRMCS	Future Railway Mobile Communication System
GNSS	Global Navigation Satellite System
GSM	Global System for Mobile communications
GSM-R	GSM Railway
HAPS	High-Altitude Platform Stations
IP	Internet Protocol
IIoT	Industrial IoT
IoT	Internet of Things
ITU	International Telecommunication Union
LEO	Low Earth Orbit
LMC	Location Management Component
LMF	Location Management Function
MAC	Medium Access Control
MCL	Maximum Coupling Loss
ML	Machine Learning
MPE	Maximum Permitted Exposure
MR	Mixed Reality
NG-RAN	Next Generation RAN
NR	New Radio
NSA	Non-Standalone
NTN	Non-Terrestrial Networks
NWDAF	NetWork Data Analytic Function
OAM	Operations And Management
PCell	Primary Cell
PDB	Packet Delay Budget
PDCCH	Physical Downlink Control Channel
PDU	Protocol Data Unit
PDCP	Packet Data Convergence Protocol
PRB	Physical Resource Block
PRS	Positioning Reference Signal
PSS	Primary Synchronization Signal
PTP	Precision Time Protocol
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying

RA	Registration Area
RAN	Radio Access Network
RedCap	Reduced Capability
RIM	Remote Interference Mitigation
RL	Reinforcement Learning
RLC	Radio Link Control
RTK	Real-Time Kinematic
RTT	Round Trip Time
RU	Radio Unit
SA	Standalone
SaaS	Synchronization as a Service
SCell	Secondary Cell
SDT	Small Data Transmission
SLA	Service Level Agreement
SPS	Semi-Persistent Scheduling
SSS	Secondary Synchronization Signal
TA	Tracking Area
TAI	TA Identifier
TCI	Transmission Configuration Indication
TDD	Time Division Duplexing
TDOA	Time Difference of Arrival
TV	Television
UAV	Uncrewed Aerial Vehicle
UE	User Equipment
UPF	User Plane Function
URSP	UE Route Selection Policy
VR	Virtual Reality
XR	eXtended Reality

About Nokia

At Nokia, we create technology that helps the world act together.

As a B2B technology innovation leader, we are pioneering the future where networks meet cloud to realize the full potential of digital in every industry.

Through networks that sense, think and act, we work with our customers and partners to create the digital services and applications of the future.

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: CID210984 (February)