

Workgroup: RAW

Published: 30 March 2020

Intended Status: Informational

Expires: 1 October 2020

Authors: J. Farkas, Ed. T. Dudda A. Shapin S. Sandberg
 Ericsson Ericsson Ericsson Ericsson

5G - Ultra-Reliable Wireless Technology with Low Latency

Abstract

This document describes the features of 5G that make it a wireless technology providing ultra-reliability, high availability, and low latency; and looks out to possibilities on the application of 5G together with IETF Deterministic Networking (DetNet).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 1 October 2020.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

- [1. Introduction](#)
- [2. Provenance and Documents](#)
- [3. General Characteristics](#)
- [4. Deployment and Spectrum](#)
- [5. Applicability to Deterministic Flows](#)
 - [5.1. System Architecture](#)
 - [5.2. Overview of The Radio Protocol Stack](#)
 - [5.3. Radio \(PHY\)](#)
 - [5.4. Scheduling and QoS \(MAC\)](#)
 - [5.5. Time-Sensitive Networking \(TSN\) Integration](#)
- [6. Summary](#)
- [7. IANA Considerations](#)
- [8. Security Considerations](#)
- [9. Acknowledgments](#)
- [10. Informative References](#)
- [Authors' Addresses](#)

1. Introduction

5G is a highly predictable scheduled wireless technology. Equipped with Ultra-Reliable Low-Latency Communication (URLLC) features, 5G provides ultra reliability and high availability as well as low latency for critical communications. That is, 5G is a Reliable Available Wireless (RAW) technology. Its characteristics make 5G perfectly suitable to be part of deterministic networks, e.g., industrial automation networks. Furthermore, 5G already includes features and capabilities for integration with deterministic wireline technologies such as IEEE 802.1 Time-Sensitive Networking (TSN) [[IEEE802.1TSN](#)] and IETF Deterministic Networking (DetNet) [[RFC8655](#)].

2. Provenance and Documents

The 3rd Generation Partnership Project (3GPP) incorporates many companies whose business is related to cellular network operation as well as network equipment and device manufacturing. All generations of 3GPP technologies provide scheduled wireless segments, primarily in licensed spectrum which is beneficial for reliability and availability.

In 2016, the 3GPP started to design New Radio (NR) technology belonging to the fifth generation (5G) of cellular networks. NR has been designed from the beginning to not only address enhanced Mobile Broadband (eMBB) services for consumer devices such as smart phones or tablets but is also tailored for future Internet of Things (IoT) communication and connected cyber-physical systems. In addition to eMBB, requirement categories have been defined on Massive Machine-Type Communication (M-MTC) for a large number of connected devices/sensors, and Ultra-Reliable Low-Latency Communication (URLLC) for connected control systems and critical communication as illustrated in [Figure 1](#). It is the URLLC capabilities that make 5G a great candidate for reliable low-latency communication. With these three corner stones, NR is a complete solution supporting the connectivity needs of consumers, enterprises, and public sector for both wide area and local area, e.g. indoor deployments. A general overview of NR can be found in [[TS38300](#)].

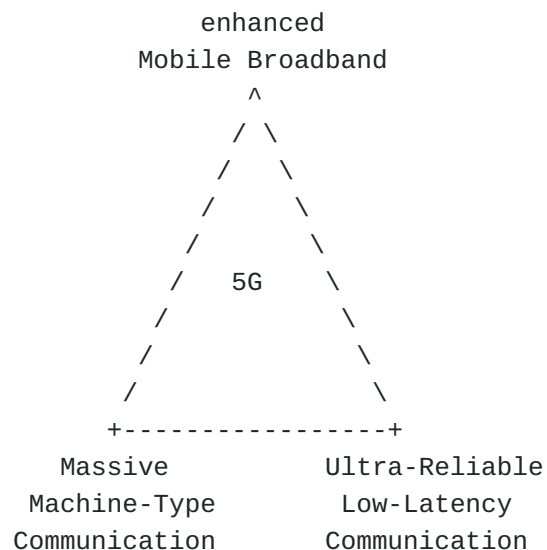


Figure 1: 5G Application Areas

As a result of releasing the first NR specification in 2018 (Release 15), it has been proven by many companies that NR is a URLLC-capable technology and can deliver data packets at 10^{-5} packet error rate within 1ms latency budget [[TR37910](#)]. Those evaluations were

consolidated and forwarded to ITU to be included in the [[IMT2020](#)] work.

In order to understand communication requirements for automation in vertical domains, 3GPP studied different use cases [[TR22804](#)] and released technical specification with reliability, availability and latency demands for a variety of applications [[TS22104](#)].

As an evolution of NR, multiple studies have been conducted in scope of 3GPP Release 16 including the following two, focusing on radio aspects:

1. Study on physical layer enhancements for NR ultra-reliable and low latency communication (URLLC) [[TR38824](#)].
2. Study on NR industrial Internet of Things (I-IoT) [[TR38825](#)].

In addition, several enhancements have been done on system architecture level which are reflected in System architecture for the 5G System (5GS) [[TS23501](#)].

3. General Characteristics

The 5G Radio Access Network (5G RAN) with its NR interface includes several features to achieve Quality of Service (QoS), such as a guaranteeably low latency or tolerable packet error rates for selected data flows. Determinism is achieved by centralized admission control and scheduling of the wireless frequency resources, which are typically licensed frequency bands assigned to a network operator.

NR enables short transmission slots in a radio subframe, which benefits low-latency applications. NR also introduces mini-slots, where prioritized transmissions can be started without waiting for slot boundaries, further reducing latency. As part of giving priority and faster radio access to URLLC traffic, NR introduces preemption where URLLC data transmission can preempt ongoing non-URLLC transmissions. Additionally, NR applies very fast processing, enabling retransmissions even within short latency bounds.

NR defines extra-robust transmission modes for increased reliability both for data and control radio channels. Reliability is further improved by various techniques, such as multi-antenna transmission, the use of multiple frequency carriers in parallel and packet duplication over independent radio links. NR also provides full mobility support, which is an important reliability aspect not only for devices that are moving, but also for devices located in a changing environment.

Network slicing is seen as one of the key features for 5G, allowing vertical industries to take advantage of 5G networks and services. Network slicing is about transforming a Public Land Mobile Network (PLMN) from a single network to a network where logical partitions are created, with appropriate network isolation, resources, optimized topology and specific configuration to serve various service requirements. An operator can configure and manage the mobile network to support various types of services enabled by 5G, for example eMBB and URLLC, depending on the different customers' needs.

Exposure of capabilities of 5G Systems to the network or applications outside the 3GPP domain have been added to Release 16 [[TS23501](#)]. Via exposure interfaces, applications can access 5G capabilities, e.g., communication service monitoring and network maintenance.

For several generations of mobile networks, 3GPP has considered how the communication system should work on a global scale with billions of users, taking into account resilience aspects, privacy regulation, protection of data, encryption, access and core network security, as well as interconnect. Security requirements evolve as demands on trustworthiness increase. For example, this has led to the introduction of enhanced privacy protection features in 5G. 5G also employs strong security algorithms, encryption of traffic, protection of signaling and protection of interfaces.

One particular strength of mobile networks is the authentication, based on well-proven algorithms and tightly coupled with a global identity management infrastructure. Since 3G, there is also mutual authentication, allowing the network to authenticate the device and the device to authenticate the network. Another strength is secure solutions for storage and distribution of keys fulfilling regulatory requirements and allowing international roaming. When connecting to 5G, the user meets the entire communication system, where security is the result of standardization, product security, deployment, operations and management as well as incident handling capabilities. The mobile networks approach the entirety in a rather coordinated fashion which is beneficial for security.

4. Deployment and Spectrum

The 5G system allows deployment in a vast spectrum range, addressing use-cases in both wide-area as well as local networks. Furthermore, 5G can be configured for public and non-public access.

When it comes to spectrum, NR allows combining the merits of many frequency bands, such as the high bandwidths in millimeter Waves (mmW) for extreme capacity locally, as well as the broad coverage

when using mid- and low frequency bands to address wide-area scenarios. URLLC is achievable in all these bands. Spectrum can be either licensed, which means that the license holder is the only authorized user of that spectrum range, or unlicensed, which means that anyone who wants to use the spectrum can do so.

A prerequisite for critical communication is performance predictability, which can be achieved by the full control of the access to the spectrum, which 5G provides. Licensed spectrum guarantees control over spectrum usage by the system, making it a preferable option for critical communication. However, unlicensed spectrum can provide an additional resource for scaling non-critical communications. While NR is initially developed for usage of licensed spectrum, the functionality to access also unlicensed spectrum was introduced in 3GPP Release 16.

Licensed spectrum dedicated to mobile communications has been allocated to mobile service providers, i.e. issued as longer-term licenses by national administrations around the world. These licenses have often been associated with coverage requirements and issued across whole countries, or in large regions. Besides this, configured as a non-public network (NPN) deployment, 5G can provide network services also to a non-operator defined organization and its premises such as a factory deployment. By this isolation, quality of service requirements, as well as security requirements can be achieved. An integration with a public network, if required, is also possible. The non-public (local) network can thus be interconnected with a public network, allowing devices to roam between the networks.

In an alternative model, some countries are now in the process of allocating parts of the 5G spectrum for local use to industries. These non-service providers then have a choice of applying for a local license themselves and operating their own network or cooperating with a public network operator or service provider.

5. Applicability to Deterministic Flows

5.1. System Architecture

The 5G system [[TS23501](#)] consists of the User Equipment (UE) at the terminal side, and the Radio Access Network (RAN) with the gNB as radio base station node, as well as the Core Network (CN). The core network is based on a service-based architecture with the central functions: Access and Mobility Management Function (AMF), Session Management Function (SMF) and User Plane Function (UPF) as illustrated in [Figure 2](#).

The gNB's main responsibility is the radio resource management, including admission control and scheduling, mobility control and radio measurement handling. The AMF handles the UE's connection status and security, while the SMF controls the UE's data sessions. The UPF handles the user plane traffic.

The SMF can instantiate various Packet Data Unit (PDU) sessions for the UE, each associated with a set of QoS flows, i.e., with different QoS profiles. Segregation of those sessions is also possible, e.g., resource isolation in the RAN and in the CN can be defined (slicing).

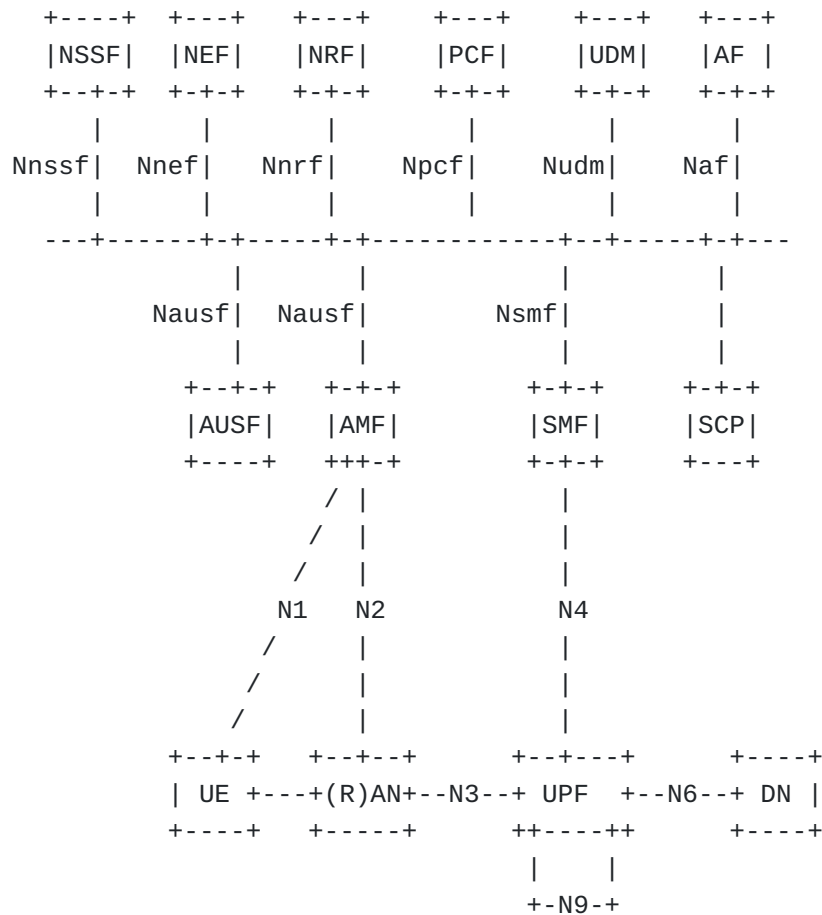


Figure 2: 5G System Architecture

To allow UE mobility across cells/gNBs, handover mechanisms are supported in NR. For an established connection, i.e., connected mode mobility, a gNB can configure a UE to report measurements of received signal strength and quality of its own and neighbouring cells, periodically or event-based. Based on these measurement reports, the gNB decides to handover a UE to another target cell/gNB. Before triggering the handover, it is hand-shaked with the target gNB based on network signalling. A handover command is then

sent to the UE and the UE switches its connection to the target cell/gNB. The Packet Data Convergence Protocol (PDCP) of the UE can be configured to avoid data loss in this procedure, i.e., handle retransmissions if needed. Data forwarding is possible between source and target gNB as well. To improve the mobility performance further, i.e., to avoid connection failures, e.g., due to too-late handovers, the mechanism of conditional handover is introduced in Release 16 specifications. Therein a conditional handover command, defining a triggering point, can be sent to the UE before UE enters a handover situation. A further improvement introduced in Release 16 is the Dual Active Protocol Stack (DAPS), where the UE maintains the connection to the source cell while connecting to the target cell. This way, potential interruptions in packet delivery can be avoided entirely.

5.2. Overview of The Radio Protocol Stack

The protocol architecture for NR consists of the L1 Physical layer (PHY) and as part of the L2, the sublayers of Medium Access Control (MAC), Radio Link Control (RLC), Packet Data Convergence Protocol (PDCP), as well as the Service Data Adaption Protocol (SDAP).

The PHY layer handles signal processing related actions, such as encoding/decoding of data and control bits, modulation, antenna precoding and mapping.

The MAC sub-layer handles multiplexing and priority handling of logical channels (associated with QoS flows) to transport blocks for PHY transmission, as well as scheduling information reporting and error correction through Hybrid Automated Repeat Request (HARQ).

The RLC sublayer handles sequence numbering of higher layer packets, retransmissions through Automated Repeat Request (ARQ), if configured, as well as segmentation and reassembly and duplicate detection.

The PDCP sublayer consists of functionalities for ciphering/deciphering, integrity protection/verification, re-ordering and in-order delivery, duplication and duplicate handling for higher layer packets, and acts as the anchor protocol to support handovers.

The SDAP sublayer provides services to map QoS flows, as established by the 5G core network, to data radio bearers (associated with logical channels), as used in the 5G RAN.

Additionally, in RAN, the Radio Resource Control (RRC) protocol, handles the access control and configuration signalling for the aforementioned protocol layers. RRC messages are considered L3 and thus transmitted also via those radio protocol layers.

To provide low latency and high reliability for one transmission link, i.e., to transport data (or control signaling) of one radio bearer via one carrier, several features have been introduced on the user plane protocols for PHY and L2, as explained in the following.

5.3. Radio (PHY)

NR is designed with native support of antenna arrays utilizing benefits from beamforming, transmissions over multiple MIMO layers and advanced receiver algorithms allowing effective interference cancellation. Those antenna techniques are the basis for high signal quality and effectiveness of spectral usage. Spatial diversity with up to 4 MIMO layers in UL and up to 8 MIMO layers in DL is supported. Together with spatial-domain multiplexing, antenna arrays can focus power in desired direction to form beams. NR supports beam management mechanisms to find the best suitable beam for UE initially and when it is moving. In addition, gNBs can coordinate their respective DL and UL transmissions over the backhaul network keeping interference reasonably low, and even make transmissions or receptions from multiple points (multi-TRP). Multi-TRP can be used for repetition of data packet in time, in frequency or over multiple MIMO layers which can improve reliability even further.

Any downlink transmission to a UE starts from resource allocation signaling over the Physical Downlink Control Channel (PDCCH). If it is successfully received, the UE will know about the scheduled transmission and may receive data over the Physical Downlink Shared Channel (PDSCH). If retransmission is required according to the HARQ scheme, a signaling of negative acknowledgement (NACK) on the Physical Uplink Control Channel (PUCCH) is involved and PDCCH together with PDSCH transmissions (possibly with additional redundancy bits) are transmitted and soft-combined with previously received bits. Otherwise, if no valid control signaling for scheduling data is received, nothing is transmitted on PUCCH (discontinuous transmission - DTX), and the base station upon detecting DTX will retransmit the initial data.

An uplink transmission normally starts from a Scheduling Request (SR) - a signaling message from the UE to the base station sent via PUCCH. Once the scheduler is informed about buffer data in UE, e.g., by SR, the UE transmits a data packet on the Physical Uplink Shared Channel (PUSCH). Pre-scheduling not relying on SR is also possible (see following section).

Since transmission of data packets require usage of control and data channels, there are several methods to maintain the needed reliability. NR uses Low Density Parity Check (LDPC) codes for data channels, Polar codes for PDCCH, as well as orthogonal sequences and Polar codes for PUCCH. For ultra-reliability of data channels, very

robust (low spectral efficiency) Modulation and Coding Scheme (MCS) tables are introduced containing very low (down to 1/20) LDPC code rates using BPSK or QPSK. Also, PDCCH and PUCCH channels support multiple code rates including very low ones for the channel robustness.

A connected UE reports downlink (DL) quality to gNB by sending Channel State Information (CSI) reports via PUCCH while uplink (UL) quality is measured directly at gNB. For both uplink and downlink, gNB selects the desired MCS number and signals it to the UE by Downlink Control Information (DCI) via PDCCH channel. For URLLC services, the UE can assist the gNB by advising that MCS targeting 10^{-5} Block Error Rate (BLER) are used. Robust link adaptation algorithms can maintain the needed level of reliability considering a given latency bound.

Low latency on the physical layer is provided by short transmission duration which is possible by using high Subcarrier Spacing (SCS) and the allocation of only one or a few Orthogonal Frequency Division Multiplexing (OFDM) symbols. For example, the shortest latency for the worst case in DL can be 0.23ms and in UL can be 0.24ms according to (section 5.7.1 in [\[TR37910\]](#)). Moreover, if the initial transmission has failed, HARQ feedback can quickly be provided and an HARQ retransmission is scheduled.

Dynamic multiplexing of data associated with different services is highly desirable for efficient use of system resources and to maximize system capacity. Assignment of resources for eMBB is usually done with regular (longer) transmission slots, which can lead to blocking of low latency services. To overcome the blocking, eMBB resources can be pre-empted and re-assigned to URLLC services. In this way, spectrally efficient assignments for eMBB can be ensured while providing flexibility required to ensure a bounded latency for URLLC services. In downlink, the gNB can notify the eMBB UE about pre-emption after it has happened, while in uplink there are two pre-emption mechanisms: special signaling to cancel eMBB transmission and URLLC dynamic power boost to suppress eMBB transmission.

5.4. Scheduling and QoS (MAC)

One integral part of the 5G system is the Quality of Service (QoS) framework [\[TS23501\]](#). QoS flows are setup by the 5G system for certain IP or Ethernet packet flows, so that packets of each flow receive the same forwarding treatment, i.e., in scheduling and admission control. QoS flows can for example be associated with different priority level, packet delay budgets and tolerable packet error rates. Since radio resources are centrally scheduled in NR,

the admission control function can ensure that only those QoS flows are admitted for which QoS targets can be reached.

NR transmissions in both UL and DL are scheduled by the gNB [TS38300]. This ensures radio resource efficiency, fairness in resource usage of the users and enables differentiated treatment of the data flows of the users according to the QoS targets of the flows. Those QoS flows are handled as data radio bearers or logical channels in NR RAN scheduling.

The gNB can dynamically assign DL and UL radio resources to users, indicating the resources as DL assignments or UL grants via control channel to the UE. Radio resources are defined as blocks of OFDM symbols in spectral domain and time domain. Different lengths are supported in time domain, i.e., (multiple) slot or mini-slot lengths. Resources of multiple frequency carriers can be aggregated and jointly scheduled to the UE.

Scheduling decisions are based, e.g., on channel quality measured on reference signals and reported by the UE (cf. periodical CSI reports for DL channel quality). The transmission reliability can be chosen in the scheduling algorithm, i.e., by link adaptation where an appropriate transmission format (e.g., robustness of modulation and coding scheme, controlled UL power) is selected for the radio channel condition of the UE. Retransmissions, based on HARQ feedback, are also controlled by the scheduler. If needed to avoid HARQ round-trip time delays, repeated transmissions can be also scheduled beforehand, to the cost of reduced spectral efficiency.

In dynamic DL scheduling, transmission can be initiated immediately when DL data becomes available in the gNB. However, for dynamic UL scheduling, when data becomes available but no UL resources are available yet, the UE indicates the need for UL resources to the gNB via a (single bit) scheduling request message in the UL control channel. When thereupon UL resources are scheduled to the UE, the UE can transmit its data and may include a buffer status report, indicating the exact amount of data per logical channel still left to be sent. More UL resources may be scheduled accordingly. To avoid the latency introduced in the scheduling request loop, UL radio resources can also be pre-scheduled.

In particular for periodical traffic patterns, the pre-scheduling can rely on the scheduling features DL Semi-Persistent Scheduling (SPS) and UL Configured Grant (CG). With these features, periodically recurring resources can be assigned in DL and UL. Multiple parallels of those configurations are supported, in order to serve multiple parallel traffic flows of the same UE.

To support QoS enforcement in the case of mixed traffic with different QoS requirements, several features have recently been introduced. This way, e.g., different periodical critical QoS flows can be served together with best effort transmissions, by the same UE. Among others, these features (partly Release 16) are: 1) UL logical channel transmission restrictions allowing to map logical channels of certain QoS only to intended UL resources of a certain frequency carrier, slot-length, or CG configuration, and 2) intra-UE pre-emption, allowing critical UL transmissions to pre-empt non-critical transmissions.

When multiple frequency carriers are aggregated, duplicate parallel transmissions can be employed (beside repeated transmissions on one carrier). This is possible in the Carrier Aggregation (CA) architecture where those carriers originate from the same gNB, or in the Dual Connectivity (DC) architecture where the carriers originate from different gNBs, i.e., the UE is connected to two gNBs in this case. In both cases, transmission reliability is improved by this means of providing frequency diversity.

In addition to licensed spectrum, a 5G system can also utilize unlicensed spectrum to offload non-critical traffic. This version of NR is called NR-U, part of 3GPP Release 16. The central scheduling approach applies also for unlicensed radio resources, but in addition also the mandatory channel access mechanisms for unlicensed spectrum, e.g., Listen Before Talk (LBT) are supported in NR-U. This way, by using NR, operators have and can control access to both licensed and unlicensed frequency resources.

5.5. Time-Sensitive Networking (TSN) Integration

The main objective of Time-Sensitive Networking (TSN) is to provide guaranteed data delivery within a guaranteed time window, i.e., bounded low latency. IEEE 802.1 TSN [[IEEE802.1TSN](#)] is a set of open standards that provide features to enable deterministic communication on standard IEEE 802.3 Ethernet [[IEEE802.3](#)]. TSN standards can be seen as a toolbox for traffic shaping, resource management, time synchronization, and reliability.

A TSN stream is a data flow between one end station (Talker) to another end station (Listener). In the centralized configuration model, TSN bridges are configured by the Central Network Controller (CNC) [[IEEE802.1Qcc](#)] to provide deterministic connectivity for the TSN stream through the network. Time-based traffic shaping provided by Scheduled Traffic [[IEEE802.1Qbv](#)] may be used to achieve bounded low latency. The TSN tool for time synchronization is the generalized Precision Time Protocol (gPTP) [[IEEE802.1AS](#)], which provides reliable time synchronization that can be used by end stations and by other TSN tools, e.g., Scheduled Traffic

[[IEEE802.1Qbv](#)]. High availability, as a result of ultra-reliability, is provided for data flows by the Frame Replication and Elimination for Reliability (FRER) [[IEEE802.1CB](#)] mechanism.

3GPP Release 16 includes integration of 5G with TSN, i.e., specifies functions for the 5G System (5GS) to deliver TSN streams such that they meet their QoS requirements. A key aspect of the integration is that the 5GS appears from the rest of the network as a set of TSN bridges, in particular, one virtual bridge per User Plane Function (UPF) on the user plane. The 5GS includes TSN Translator (TT) functionality for the adaptation of the 5GS to the TSN bridged network and for hiding the 5GS internal procedures. The 5GS provides the following components:

1. interface to TSN controller, as per [[IEEE802.1Qcc](#)] for the fully centralized configuration model
2. time synchronization via reception and transmission of gPTP PDUs [[IEEE802.1AS](#)]
3. low latency, hence, can be integrated with Scheduled Traffic [[IEEE802.1Qbv](#)]
4. reliability, hence, can be integrated with FRER [[IEEE802.1CB](#)]

[Figure 2](#) shows an illustration of 5G-TSN integration where an industrial controller (Ind Ctrlr) is connected to industrial Input/Output devices (I/O dev) via 5G. The 5GS can directly transport Ethernet frames since Release 15, thus, end-to-end Ethernet connectivity is provided. The 5GS implements the required interfaces towards the TSN controller functions such as the CNC, thus adapts to the settings of the TSN network. A 5G user plane virtual bridge interconnects TSN bridges or connects end stations, e.g., I/O devices to the network. Note that the introduction of 5G brings flexibility in various aspects, e.g., more flexible network topology because a wireless hop can replace several wireline hops thus significantly reduce the number of hops end-to-end. [[ETR5GTSN](#)] dives more into the integration of 5G with TSN.

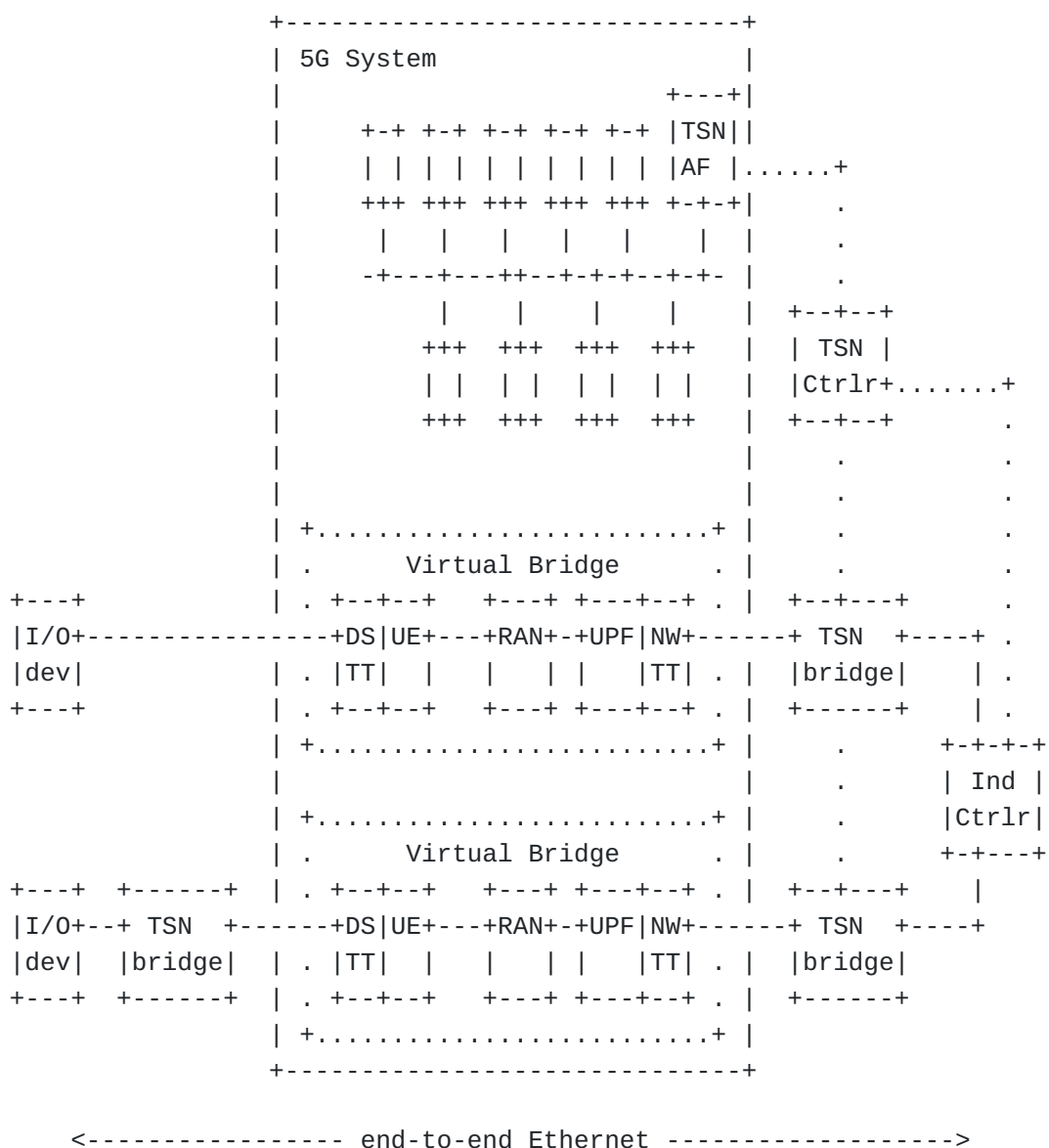


Figure 3: 5G - TSN Integration

NR supports accurate reference time synchronization in 1us accuracy level. Since NR is a scheduled system, an NR UE and a gNB are tightly synchronized to their OFDM symbol structures. A 5G internal reference time can be provided to the UE via broadcast or unicast signaling, associating a known OFDM symbol to this reference clock. The 5G internal reference time can be shared within the 5G network, i.e., radio and core network components. For the interworking with gPTP for multiple time domains, the 5GS acts as a virtual gPTP time-aware system and supports the forwarding of gPTP time synchronization information between end stations and bridges through the 5G user plane TTs. These account for the residence time of the 5GS in the time synchronization procedure. One special option is when the 5GS internal reference time is not only used within the

5GS, but also to the rest of the devices in the deployment, including connected TSN bridges and end stations.

Redundancy architectures were specified in order to provide reliability against any kind of failure on the radio link or nodes in the RAN and the core network, Redundant user plane paths can be provided based on the dual connectivity architecture, where the UE sets up two PDU sessions towards the same data network, and the 5G system makes the paths of the two PDU sessions independent as illustrated in [Figure 5](#). There are two PDU sessions involved in the solution: the first spans from the UE via gNB1 to UPF1, acting as the first PDU session anchor, while the second spans from the UE via gNB2 to UPF2, acting as second the PDU session anchor. The independent paths may continue beyond the 3GPP network. Redundancy Handling Functions (RHF) are deployed outside of the 5GS, i.e., in Host A (the device) and in Host B (the network). RHF can implement replication and elimination functions as per [[IEEE802.1CB](#)] or the Packet Replication, Elimination, and Ordering Functions (PREOF) of IETF Deterministic Networking (DetNet) [[RFC8655](#)].

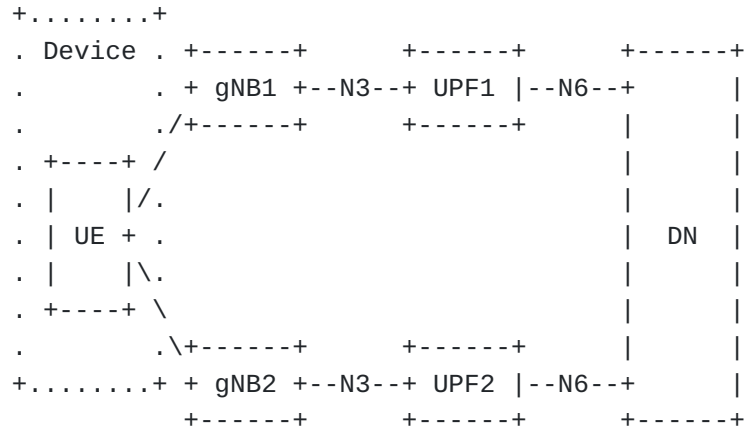


Figure 4: Reliability with Single UE

An alternative solution is that multiple UEs per device are used for user plane redundancy as illustrated in [Figure 5](#). Each UE sets up a PDU session. The 5GS ensures that those PDU sessions of the different UEs are handled independently internal to the 5GS. There is no single point of failure in this solution, which also includes RHF outside of the 5G system, e.g., as per FRER or as PREOF specifications.

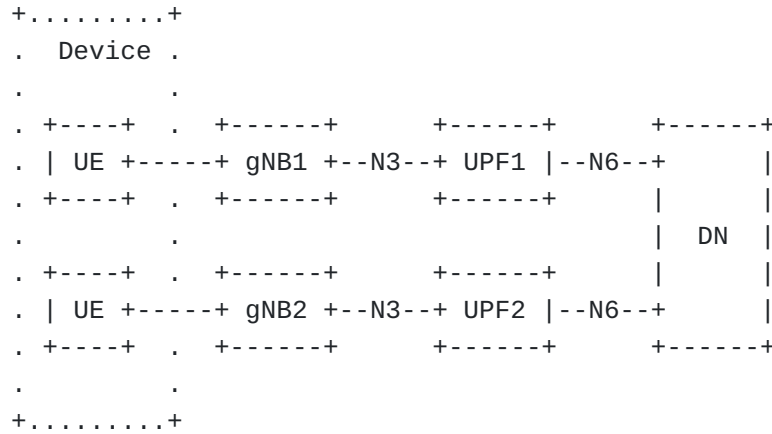


Figure 5: Reliability with Dual UE

Note that the abstraction provided by the RHF and the location of the RHF being outside of the 5G system make 5G equally supporting integration for reliability both with FRER of TSN and PREOF of DetNet as they both rely on the same concept.

Note also that TSN is the primary subnetwork technology for DetNet. Thus, the DetNet over TSN work, e.g., [[I-D.ietf-detnet-ip-over-tsn](#)], can be leveraged via the TSN support built in 5G.

6. Summary

5G technology enables deterministic communication. Based on the centralized admission control and the scheduling of the wireless resources, licensed or unlicensed, quality of service such as latency and reliability can be guaranteed. 5G contains several features to achieve ultra-reliable and low latency performance, e.g., support for different OFDM numerologies and slot-durations, as well as fast processing capabilities and redundancy techniques that lead to achievable latency numbers of below 1ms with reliability guarantees up to 99.999%.

5G also includes features to support Industrial IoT use cases, e.g., via the integration of 5G with TSN. This includes 5G capabilities for each TSN component, latency, resource management, time synchronization, and reliability. Furthermore, 5G support for TSN can be leveraged when 5G is used as subnet technology for DetNet, in combination with or instead of TSN, which is the primary subnet for DetNet. In addition, the support for integration with TSN reliability was added to 5G by making DetNet reliability also applicable, thus making 5G DetNet ready. Moreover, providing IP service is native to 5G.

Overall, 5G provides scheduled wireless segments with high reliability and availability. In addition, 5G includes capabilities for integration to IP networks.

7. IANA Considerations

This document does not require IANA action.

8. Security Considerations

5G includes security mechanisms as defined by 3GPP.

9. Acknowledgments

The authors acknowledge the work of all from Ericsson Research who contributed to the subject in any form.

10. Informative References

- [TR37910] "3GPP TR 37.910, Study on self evaluation towards IMT-2020 submission", <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3190>>.
- [TR38824] "3GPP TR 38.824, Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC)", <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3498>>.
- [TR38825] "3GPP TR 38.825, Study on NR industrial Internet of Things (IoT)", <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3492>>.
- [TS22104] "3GPP TS 22.104, Service requirements for cyber-physical control applications in vertical domains", <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3528>>.
- [TR22804] "3GPP TR 22.804, Study on Communication for Automation in Vertical domains (CAV)", <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3187>>.
- [TS23501] "3GPP TS 23.501, System architecture for the 5G System (5GS)", <<https://portal.3gpp.org/desktopmodules/>>.

[Specifications/SpecificationDetails.aspx?specificationId=3144](#)>.

[TS38300] "3GPP TS 38.300, NR Overall description", <<https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3191>>.

[IMT2020] "ITU towards IMT for 2020 and beyond", <<https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2020/Pages/default.aspx>>.

[RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", RFC 8655, DOI 10.17487/RFC8655, October 2019, <<https://www.rfc-editor.org/info/rfc8655>>.

[I-D.ietf-detnet-ip-over-tsn]

Varga, B., Farkas, J., Malis, A., and S. Bryant, "DetNet Data Plane: IP over IEEE 802.1 Time Sensitive Networking (TSN)", Work in Progress, Internet-Draft, draft-ietf-detnet-ip-over-tsn-02, 6 March 2020, <<https://tools.ietf.org/html/draft-ietf-detnet-ip-over-tsn-02>>.

[IEEE802.1TSN] IEEE 802.1, "Time-Sensitive Networking (TSN) Task Group", <<http://www.ieee802.org/1/pages/tsn.html>>.

[IEEE802.1AS] IEEE, "IEEE Standard for Local and metropolitan area networks -- Timing and Synchronization for Time-Sensitive Applications", IEEE 802.1AS-2020, <https://standards.ieee.org/content/ieee-standards/en/standard/802_1AS-2020.html>.

[IEEE802.1CB] IEEE, "IEEE Standard for Local and metropolitan area networks -- Frame Replication and Elimination for Reliability", DOI 10.1109/IEEESTD.2017.8091139, IEEE 802.1CB-2017, , <<https://ieeexplore.ieee.org/document/8091139>>.

[IEEE802.1Qbv] IEEE, "IEEE Standard for Local and metropolitan area networks -- Bridges and Bridged Networks -- Amendment 25: Enhancements for Scheduled Traffic", IEEE 802.1Qbv-2015, <<https://ieeexplore.ieee.org/document/7440741>>.

[IEEE802.1Qcc] IEEE, "IEEE Standard for Local and metropolitan area networks -- Bridges and Bridged Networks -- Amendment 31: Stream Reservation Protocol (SRP) Enhancements and

Performance Improvements", IEEE 802.1Qcc-2018, <<https://ieeexplore.ieee.org/document/8514112>>.

[**IEEE802.3**] IEEE, "IEEE Standard for Ethernet", IEEE 802.3-2018, <<https://ieeexplore.ieee.org/document/8457469>>.

[**ETR5GTSN**] Farkas, J., Varga, B., Miklos, G., and J. Sachs, "5G-TSN integration meets networking requirements for industrial automation", Ericsson Technology Review, Volume 9, No 7, August 2019, <<https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/5g-tsn-integration-for-industrial-automation>>.

Authors' Addresses

Janos Farkas (editor)
Ericsson
Budapest
Magyar tudosok korutja 11
1117
Hungary

Email: janos.farkas@ericsson.com

Torsten Dudda
Ericsson
Ericsson Allee 1
52134 Herzogenrath
Germany

Email: torsten.dudda@ericsson.com

Alexey Shapin
Ericsson
Laboratoriegrand 11
977 53 Lulea
Sweden

Email: alexey.shapin@ericsson.com

Sara Sandberg
Ericsson
Laboratoriegrand 11
977 53 Lulea
Sweden

Email: sara.sandberg@ericsson.com