Ultra Reliable Low Latency Communication for 5G New Radio

IEEE Workshop on 5G Technologies for Tactical and First Responder Networks

Rapeepat Ratasuk, Principal Research Specialist
Nokia Bell Labs
October 23rd, 2018
5G and URLLC

Enhanced Mobile Broadband (eMBB)

Ultra Reliable Low Latency communication

Ultra Reliable
< 10^{-3} outage

Zero mobility interruption

Unlimited experience

>10 Gbps peak data rates

<4 ms radio latency

100 Mbps whenever needed

10,000 x more traffic

Ultra Low Cost

<1 ms radio latency

Massive machine type communication

1,000,000 devices per km²

164 dB MCL

>15 years on battery

mMTC ultra low cost

Range

Zero mobility interruption

Ultra Reliable

Massive machine type communication

164 dB MCL

>15 years on battery

mMTC ultra low cost

URLLC is one of the three usage scenarios identified by ITU IMT-2020
# Use Cases

## Selected URLLC Use Cases

**Factory Automation (local area)**
- **Motion Control**
  - Indoor
  - Most stringent availability, latency, and jitter
  - Limited mobility
- **Mobile Robots/AGVs**
  - Indoor & outdoor
  - Stringent availability, latency, and jitter
  - Ultra reliable mobility (following pre-defined paths)

**Harbor Automation**
- Remote control of cranes and vehicles
- Indoor & outdoor
- Mixed data traffic types
- High reliability

**Smart City Automation (wide area)**
- **Smart Grid Protection and Control**
  - Outdoor
  - Stringent requirements on availability, latency, and jitter
- **Autonomous Vehicles**
  - Outdoor
  - Stringent requirements on availability, latency, and jitter
  - Ultra reliable mobility (over wide area)

---

AGV: Automatic guided vehicle
# URLLC Use Cases and Requirements (TS 22.261)

Table 7.2.2-1 Performance requirements for low-latency and high-reliability scenarios.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>End-to-end latency</th>
<th>Communication service availability</th>
<th>Reliability</th>
<th>User experienced data rate</th>
<th>Payload size</th>
<th>Traffic density</th>
<th>Connection density</th>
<th>Service area dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discrete automation – motion control</td>
<td>1 ms</td>
<td>99,9999%</td>
<td>99,9999%</td>
<td>1 Mbps up to 10 Mbps</td>
<td>Small</td>
<td>1 Tbps/km²</td>
<td>100 000/km²</td>
<td>100 x 100 x 30 m</td>
</tr>
<tr>
<td>Discrete automation</td>
<td>10 ms</td>
<td>99,9%</td>
<td>99,9%</td>
<td>10 Mbps</td>
<td>Small to big</td>
<td>1 Tbps/km²</td>
<td>100 000/km²</td>
<td>1000 x 1000 x 30 m</td>
</tr>
<tr>
<td>Process automation – remote control</td>
<td>50 ms</td>
<td>99,9999%</td>
<td>99,9999%</td>
<td>1 Mbps up to 100 Mbps</td>
<td>Small to big</td>
<td>100 Gbps/km²</td>
<td>1 000/km²</td>
<td>300 x 300 x 50 m</td>
</tr>
<tr>
<td>Process automation – monitoring</td>
<td>50 ms</td>
<td>99,9%</td>
<td>99,9%</td>
<td>1 Mbps</td>
<td>Small</td>
<td>10 Gbps/km²</td>
<td>10 000/km²</td>
<td>300 x 300 x 50 m</td>
</tr>
<tr>
<td>Electricity distribution – medium voltage</td>
<td>25 ms</td>
<td>99,9%</td>
<td>99,9%</td>
<td>10 Mbps</td>
<td>Small to big</td>
<td>10 Gbps/km²</td>
<td>1 000/km²</td>
<td>100 km along power line</td>
</tr>
<tr>
<td>Electricity distribution – high voltage</td>
<td>5 ms</td>
<td>99,9999%</td>
<td>99,9999%</td>
<td>10 Mbps</td>
<td>Small</td>
<td>100 Gbps/km²</td>
<td>1 000/km²</td>
<td>200 km along power line</td>
</tr>
<tr>
<td>Intelligent transport systems – infrastructure backhaul</td>
<td>10 ms</td>
<td>99,9999%</td>
<td>99,9999%</td>
<td>10 Mbps</td>
<td>Small to big</td>
<td>10 Gbps/km²</td>
<td>1 000/km²</td>
<td>2 km along a road</td>
</tr>
<tr>
<td>Tactile interaction</td>
<td>0,5 ms</td>
<td>[99,999%]</td>
<td>[99,999%]</td>
<td>[Low]</td>
<td>[Small]</td>
<td>[Low]</td>
<td>[Low]</td>
<td>TBC</td>
</tr>
<tr>
<td>Remote control</td>
<td>[5 ms]</td>
<td>[99,999%]</td>
<td>[99,999%]</td>
<td>[From low to 10 Mbps]</td>
<td>[Small to big]</td>
<td>[Low]</td>
<td>[Low]</td>
<td>TBC</td>
</tr>
</tbody>
</table>

Various requirements from different URLLC services
### URLLC Requirements from 3GPP RAN TR 38.913 (v14.3.0, August 2017)

<table>
<thead>
<tr>
<th>Requirements</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>7.4 Control plane latency</strong></td>
<td>10ms</td>
</tr>
<tr>
<td>From a battery efficient state to start of continuous data transfer</td>
<td></td>
</tr>
<tr>
<td><strong>7.5 User plane latency</strong></td>
<td>0.5ms for both UL and DL</td>
</tr>
<tr>
<td>The time it takes to successfully deliver an application layer packet/message from the radio protocol layer 2/3 SDU ingress point to the radio protocol layer 2/3 SDU egress point via the radio interface in both uplink and downlink directions</td>
<td></td>
</tr>
<tr>
<td><strong>7.7 Mobility interruption time</strong></td>
<td>0ms</td>
</tr>
<tr>
<td>The shortest time duration supported by the system during which a user terminal cannot exchange user plane packets.</td>
<td></td>
</tr>
<tr>
<td><strong>7.9 Reliability</strong></td>
<td>1-10⁻⁵ for 32 bytes with a user plane latency of 1ms</td>
</tr>
<tr>
<td>Reliability is defined as the success probability R of transmitting X bits within a certain delay at a certain channel quality (e.g. coverage-edge).</td>
<td></td>
</tr>
</tbody>
</table>
Flexible NR Framework

- NR provides flexible framework to support different services and QoS requirements
  - Scalable slot duration, mini-slot and slot aggregation
  - Self-contained slot structure
  - Traffic preemption for URLLC
  - Support for different numerologies for different services

- NR transmission is well-contained in time and frequency
  - Future features can be easily accommodated

© Nokia 2018
Ultra Reliable and Low Latency Communication
It is not only about radio…

- **Hybrid connectivity** (e.g. multi-operator, multi-RAT)
- **5GC concepts** (e.g. admission & policy ctrl, flow based QoS, context awareness, CP decomposition)
- **NR L2-3 concepts** (e.g. multi-connectivity, conditional handover, SON-MRO)
- **NR L1 concepts** (e.g. shorter TTI, multiplexing, grant based DL, grant free UL, joint transmission, control channel enhancements)
- **O&M** (e.g. E2E reliability monitoring and verification, fault detection/prevention)
- **NW planning & optimization** (e.g. channel prediction, deployment self-optimization)
- **Flexible architecture** (e.g. lower latency with MEC, SDN & VNF reliability)
- **NW Slicing** (e.g. slice provisioning, dynamic resource usage)
- **Techno-economics** (e.g. TCO analysis, e2e reliability modeling, reliability cost functions)
- **Security** (e.g. protection against cyber attacks)
- **Transport** (e.g. rapid rerouting, PON slicing and scheduling)

E2E Reliability
Timeline – URLLC Related Items

- **2017**
  - Q4
    - Rel-15 NSA (option-3) freeze
    - Rel-15 SA (option-2) freeze

- **2018**
  - Q1
    - Rel-15 NSA (option-3) ASN.1
  - Q2
    - Rel-15 SA (option-2) ASN.1
  - Q3
    - Rel-15 late drop freeze
  - Q4
    - Rel-15 late drop ASN.1

- **2019**
  - Q1
    - Rel-16 NR URLLC L1 Enhancements SI RP-181477
  - Q2
    - Rel-16 NR URLLC L1 Enhancements WI
  - Q3
    - Rel-16 NR Industrial IoT SI RP-181479
  - Q4
    - Rel-16 late freeze
    - Rel-15 SA (option-2) ASN.1

- **2020**
  - Q1
    - Rel-16 late ASN.1

We are here
Physical Channels & Physical Signals

**PDSCH**
DL shared channel

**PBCH**
Broadcast channel

**PDCCH**
DL control channel

**DL Physical Signals**
- Demodulation Ref (DM-RS)
- Phase-tracking Ref (PT-RS)
- Ch State Inf Ref (CSI-RS)
- Primary Sync (PSS)
- Secondary Sync (SSS)

**gNodeB**

**PUSCH**
UL shared channel

**PUCCCH**
UL control channel

**PRACH**
Random access channel

**User Equipment**

**UL Physical Signals**
- Demodulation Ref (DM-RS)
- Phase-tracking Ref (PT-RS)
- Sounding Ref (SRS)
Scalable NR Numerology

- NR supports scalable numerology to address different spectrum, bandwidth, deployment and services.
- Sub-carrier spacing (SCS) of 15, 30, 60, 120 kHz is supported for data channels.
- $2^n$ scaling of SCS allows for efficient FFT processing.

### Supported data numerologies in Frequency Range 1 (FR1) (below 6 GHz):

<table>
<thead>
<tr>
<th>Carrier bandwidth (MHz)</th>
<th>SCS (kHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>10</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>15</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>20</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>25</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>30</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>40</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>50</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>60</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>80</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>90</td>
<td>15, 30, 60</td>
</tr>
<tr>
<td>100</td>
<td>15, 30, 60</td>
</tr>
</tbody>
</table>

### Supported data numerologies in FR2 (above 24 GHz):

<table>
<thead>
<tr>
<th>Carrier BW (MHz)</th>
<th>SCS (kHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>60, 120</td>
</tr>
<tr>
<td>100</td>
<td>60, 120</td>
</tr>
<tr>
<td>200</td>
<td>60, 120</td>
</tr>
<tr>
<td>400</td>
<td>60, 120</td>
</tr>
</tbody>
</table>
Scalable NR Slot Duration

- One slot is comprised of 14 symbols
  - Slot length depends on SCS – 1ms for 15 kHz SCS to 0.125ms for 120 kHz SCS
- Mini-slot (2, 4, or 7 symbols) can be allocated for shorter transmissions
- Slots can also be aggregated for longer transmissions
- A frame (10 ms) and sub-frame (1 ms) provide an SCS-agnostic time reference
3GPP Rel-15 URLLC Toolbox

Low latency

New numerology/slot/mini-slot/bi-directional structure
Non-slot (i.e. mini-slot) based scheduling
Downlink multiplexing between URLLC and eMBB services: pre-emptive scheduling
Uplink grant free transmission
Reduced processing time
Short PUCCH (one or two symbols)

Reliability

Blind repetitions (i.e. K-repetition)
Micro-diversity (e.g. Rank-1 MIMO)
Slot-aggregation for PDSCH and PUSCH
Multi-slot PUCCH
Configurable BLER targets for CQI report
URLLC MCS/CQI table
PDCP layer data duplication: allows a packet to be transmitted on two different carriers
Mini-slot Scheduling (Non-slot based Scheduling)

- **Main use cases:**
  - **Low latency** when using small SCS (e.g. 15 kHz)
  - Supporting **TDM within slot** when operating at high carrier frequencies and with RF beamforming

- **Design principles for mini-slot:**
  - Mini-slot is a shortened version of the slot
    - *Common DMRS structures between slot and mini-slot*
    - *Common control channel structures*
  - Flexible HARQ/scheduling timing with symbol level granularity
  - Supported mini-slot lengths
    - 2, 4, 7 OFDM symbols will have high priority
    - Other lengths can be considered later

Mini-slot plays a key role for latency reduction especially for smaller subcarrier spacing.
Example Latency Analysis with Mini-Slot Scheduling

Data arrives

DL

OFDM Symbol

Scheduling grant

DCI

PDSCH data transmission (2 symbols)

UE processing time

UL

gNB processing time and waiting for PDCCH monitoring occasions

PUCCH (1 symbol)

A/N

gNB processing time

Scheduling grant for re-transmission

PDSCH data transmission (2 symbols)
UL Grant-free Transmission

- 5G network can configure dedicated SPS resource to one UE or a shared SPS resource allocation pattern to a group of user devices with similar characteristics.
- In case of **dedicated resource per URLLC UE**, no collision and similar operation as in LTE.
- In case of shared SPS resource:
  - Multiple users share the same SPS resource allocation pattern and use **contention based access** with load control at base station.
UL Grant-free Transmission Types

- Two types of operation
  - Type 1: only based on RRC (re)configuration without any L1 signaling (released also by RRC).
  - Type 2: based on both RRC configuration and L1 signaling for activation/deactivation, LTE SPS-alike

UEs are pre-allocated with resources and other necessary parameters before grant-free transmission

Using MAC CE to acknowledge PHY control → minimizing misunderstanding on the configuration.

UL grant-free transmission reducing overall latency by removing the scheduling process
DL Multiplexing: Pre-emptive Scheduling

UE #1 (e.g. with eMBB traffic) is scheduled on the shared channel with a certain TTI size.

Latency critical data arrives for UE #2, and is immediately scheduled by puncturing the ongoing transmission to UE #1, rather than waiting for its completion, i.e. avoiding additional latency.

UE #1 receives the scheduled transmission, where part of it is punctured.

Effective scheduling to reduce DL latency
Enhancements for latency and reliability in Radio and E2E
Rel.16 URLLC aims for higher reliability and better efficiency

- K-repetition
- PDCP duplication
- Various “baseline” features (e.g. CQI/MCS tables for URLLC, Rank-1 MIMO)
- Flexible architecture
- Short TTI/non-slot based scheduling
- UL grant free
- DL pre-emptive scheduling
- Reduced processing time
- Short PUCCH

- High Reliability (>99.9999%)
- Multi-connectivity enhancements
- Multi-TRP
- 0ms handover interruption
- Various L1 enhancements (PDCCH, UCI, PUSCH, processing time line, ...)
- UL GF enhancements
- QoS monitoring support for URLLC

- Low Latency (<1ms)
- Data duplication enhancements
- Scheduling enhancements for intra- and inter-UE mux of eMBB and URLLC
- ...

- URLLC Efficiency [bps/Hz]
- R15 use case improvements (e.g. AR/VR)
- New use cases with higher requirements (e.g. factory automation, transport industry, electrical power distribution)

Key Rel.16 SI/WIs

<table>
<thead>
<tr>
<th>SI/WIs</th>
<th>Primary WG</th>
</tr>
</thead>
<tbody>
<tr>
<td>FS_5G_URLLC (SP-180118)</td>
<td>SA2</td>
</tr>
<tr>
<td>NR_IIoT (RP-181479)</td>
<td>RAN2</td>
</tr>
<tr>
<td>NR_Mob_enh (RP-181433)</td>
<td>RAN2</td>
</tr>
<tr>
<td>NR_eURLLC_L1 (RP-181477)</td>
<td>RAN1</td>
</tr>
<tr>
<td>NR eMIMO (RP-181453)</td>
<td>RAN1</td>
</tr>
</tbody>
</table>

Contacts: J. Moilanen, K. Pedersen, Zexian Li
Industrial IoT with 5G and NR

5G Time Sensitive Communications (TSC)
Support for Wireless Industrial Ethernet (IEEE TSN) and Multi-hop/Internet deterministic communications (aka IETF DetNet)

New R16 enablers on top of enhanced URLLC

i. Common understanding of global time among devices and network

ii. Network to support bounded latencies and error rates for packet transport (both min and max latency)

iii. More detailed service descriptors for TSC flows (condition for deterministic forwarding in TSN/DetNet)

Key Rel.16 SI/WIs

<table>
<thead>
<tr>
<th>Key Rel.16 SI/WIs</th>
<th>Primary WG</th>
</tr>
</thead>
<tbody>
<tr>
<td>NR_IIoT (RP-181479)</td>
<td>RAN2</td>
</tr>
<tr>
<td>FS_Vertical_LAN (SP-180507)</td>
<td>SA2</td>
</tr>
</tbody>
</table>

[i, ii] Enhanced URLLC and timing/gating flow mechanisms, optimization for Ethernet/TSN flows

[ii, iii] Absolute E2E time synchronization to grand master clock (5GS internal or external) and sync capability for connected NEs via gPTP (both UE and UPF side)

© Nokia 2018
Summary

• Diverse requirements expected from different URLLC use cases.

• 5G URLLC is one of the strong enablers of expansion to newly emerging vertical markets.

• Basic URLLC framework has been specified in NR Rel-15. In the studied scenario, URLLC requirement of $10^{-5}$ within 1ms can be achieved.

• The coming Rel-16 will further enhance URLLC features especially e.g. supporting industrial wireless Ethernet.