

# Proactive RAN Resource Reservation for URLLC Vehicular Slice

Nathalie NADDEH<sup>1,3</sup>, Sana BEN JEMAA<sup>1</sup>, Salah Eddine ELAYOUBI<sup>2</sup>, Tijani CHAHED<sup>3</sup>

<sup>1</sup>Orange Labs; Chatillon, France

<sup>2</sup>Université Paris Saclay; CentraleSupélec; L2S, CNRS, Gif-Sur-Yvette, France

<sup>3</sup>Institut Polytechnique de Paris; Telecom SudParis; Palaiseau, France

**Abstract**—Ultra-Reliable Low Latency Communications (URLLC) is a key service in fifth generation (5G) networks, that requires stringent Quality of Service (QoS) in terms of latency and reliability. As URLLC services may require specific numerology and/or specific channel access and re-transmission strategies, network slicing has been proposed as a solution for multiplexing them with other services such as enhanced Mobile Broadband (eMBB). Once the URLLC slice is configured and resources are dimensioned and allocated to it, URLLC performance targets should be attained thanks to the 5G New Radio (NR) low latency and high reliability features. However, in vehicular services such as safety message exchange, URLLC slice resource dimensioning cannot be static due to the varying number of vehicles in the cell. We show in this paper how the delay for slice reconfiguration alters the URLLC performance and propose a proactive resource reservation scheme that anticipates slice needs and allows ensuring URLLC targets. In order to reduce the impact of this proactive reservation on eMBB performance, we make use of vehicle trajectory prediction and show that limiting anticipated reservation to fewer cells allows reaching the target URLLC QoS with a limited degradation of the network capacity.

**Index Terms**—5G, RAN Slicing, Vehicular URLLC, Resource Reservation.

## I. INTRODUCTION

The fifth Generation New Radio (5G-NR) is being designed to support different types of services: enhanced Mobile Broad Band (eMBB) demanding high data rates (up to 1 Gbps), Ultra-Reliable Low Latency Communication (URLLC) requiring a very high reliability (99.9999%) and low latency (1 ms) and massive Machine Type Communication (mMTC) requiring higher connectivity (up to 1 Million connections/km<sup>2</sup>) [1]. An overview of these 5G communication services can be found in [2] and [3].

For the URLLC service, 5G-NR has defined several enabling technologies such as mini-slot, grant-free and semi-persistent scheduling [4]. As these features are URLLC-specific, network slicing has been defined as a solution enabling the accommodation of different types of services on the same infrastructure; each slice is designed in such a way so as to respect the performance requirements committed with the corresponding Verticals in the Service Level Agreement (SLA). However, satisfying jointly these requirements can be challenging, especially in dynamic environments, and in the presence of other types of traffic, such as eMBB.

Several works discussed the efficient multiplexing of URLLC and eMBB slices. In [5], the authors present a new

scheduling algorithm allowing a joint eMBB and URLLC scheduling process. URLLC traffic is dynamically multiplexed through puncturing/superposition of eMBB traffic. The resource allocation algorithm presented in [6] targets maximizing resources utilization and increasing eMBB throughput, but fails to attain low Vehicle-to-everything (V2X) latency. The authors in [7] propose a priority-based resource reservation mechanism that aims to decrease URLLC delay and increase reliability, but their solution does not reach the 99,9999% and has a latency of 10ms.

Also in [8], the puncturing mechanism is used on eMBB traffic and a new recovery technique for lost eMBB packets for efficient transmission is applied. The results show that puncturing or preemptive scheduling is efficient in terms of eMBB throughput and URLLC latency. These results prove that URLLC performance can be achieved, provided that a proper slice resource dimensioning is performed. Otherwise, eMBB resource preemption is possible in the down-link, in cases where eMBB resources are compliant with URLLC configuration. For the up-link and for services that require specific numerology (e.g. vehicular services that need to combat Doppler effect), on-the-fly resource preemption is not possible and a cell reconfiguration is needed. Such a reconfiguration may require a Radio Resource Control (RRC) reconfiguration procedure [9] and may involve Bandwidth Part (BWP) reconfiguration that takes 80 to 100 ms [10]. Such a delay is not acceptable for URLLC, and leads to a burst of packet losses (due to exceeded delay), each time a vehicle joins a new cell that has not originally been properly configured with the required URLLC slice resources.

In this paper, we propose a proactive resource allocation and slice configuration approach for URLLC vehicular services. Our baseline scheme is a reactive one, where resources are allocated upon user arrival to the cell, leading to packet losses during the cell reconfiguration procedure. We then propose three flavors of proactive allocation: a generalized, static resource reservation scheme where the maximal amount of needed resources is reserved for URLLC on all cells, an anticipation scheme where only the neighboring cells are reconfigured, and a trajectory prediction scheme that reserves the resources on the predicted trajectory of vehicles. We show that the exploitation of the knowledge of neighboring cells in the resource allocation procedure allows achieving the URLLC performance with a lower impact on eMBB

performance compared to the generalized, static reservation approach. The more sophisticated trajectory prediction scheme leads to an even lesser impact on eMBB, at the cost of higher implementation complexity.

The rest of the paper is organized as follows: we describe in section II the system model and quantify the impact of slice reconfiguration delay on the URLLC performance. We detail in section III the three proactive resource allocation scenarios. Simulation results and discussion are shown in section IV. Section V concludes the paper and indicates some future work perspectives.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. Vehicular service slice description

The URLLC class of services is intended to critical scenarios where a latency of few milliseconds or a very small packet loss can cause serious perturbation to the end application. In the vehicular scenarios context, URLLC is conveyed through V2X technologies, and has to cover several applications such as autonomous driving, vehicle platooning, on-board mission critical applications like in connected ambulances, etc. Their QoS requirements in terms of latency and reliability are very stringent compared to eMBB, with a latency target less than 1 ms and reliability equal to 99.9999%. This requires a specific numerology, with a small Transmission Time Interval (TTI) and a Doppler-proof design, robust channel coding and adapted channel access with replication in a contention-based manner. This numerology is to be selected based on an algorithm that takes as input the average Signal to Noise Ratio (SNR), the Doppler and delay spread [11]. One or more specific slices have thus to be dedicated to the vehicular services, with adequately configured numerology in a specific bandwidth part.

In this work, we consider two slices: V2X URLLC and eMBB. Each slice has a reserved amount of Physical Resource Blocks (PRBs) based on SLA requirements and objectives: packet loss on the order of  $10^{-6}$  and critical latency of 1ms for URLLC, along with high throughput for eMBB.

### B. Impact of slice reconfiguration delay

When an URLLC user arrives in a cell, if the resources corresponding to its slice are available, they can be instantaneously allocated to him as enabled by the seamless handover enhancements of 5G NR [12]. However, if the resources of the slice are not sufficient to serve the new user and the existing ones, a reconfiguration of the cell has to be performed so that the numerology on a part of the physical resources is changed. This reconfiguration requires RRC procedures, that may take up to 80-100 ms [9]. During this reconfiguration time, arriving packets are dropped, which affects negatively URLLC reliability performance<sup>1</sup>.

<sup>1</sup>Note that downlink preemption is traditionally regarded as a solution for ensuring URLLC QoS without resource reservation, but this is possible only when the resources for eMBB and URLLC slices are configured with the same numerology (subcarrier spacing, Cyclic Prefix (CP), channel access), but a different minislot size. If not, RRC reconfiguration is needed before reusing eMBB resources by URLLC.

In order to capture the impact of this slice reconfiguration delay on the URLLC performance, we implemented the vehicular slice in a simulator, whose complete description is given in Section IV. We illustrate in Figure 1 the evolution of the URLLC packet loss during the simulation time. In this simulation, a minimal amount of resources is allocated for URLLC, and the reservation is increased when new vehicles join the cell. We observe that, with the mobility of vehicles leading to handovers between cells, reliability reaches unacceptable levels in steady state. The packet loss rate decreases when simulation time advances, as a subset of cells will be reconfigured with an amount of resources that may be sufficient for serving the incoming vehicles, but even in this case, packet loss is at an unacceptable QoS level for URLLC, around  $3 \times 10^{-5}$ .

Reliability degrades even further as the speed of vehicles increases, as illustrated in Figure 2, as the reconfiguration time becomes more significant with respect to the sojourn time of the vehicle in the cell. These results call for proactive reservation schemes, where slice reconfiguration occurs before the vehicle enters the cell.

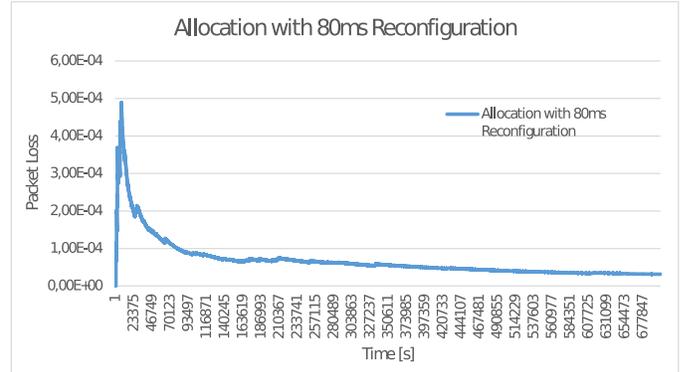


Fig. 1. URLLC packet loss illustration

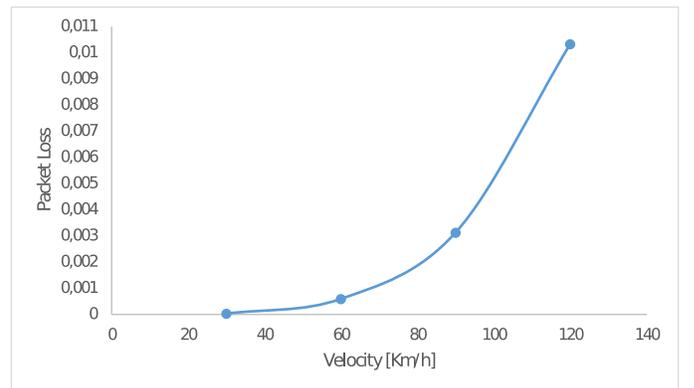


Fig. 2. Impact of vehicle velocity on reliability

## III. PROACTIVE RESOURCE RESERVATION SCHEMES

Based on the observation of degraded URLLC performance due to the slice reconfiguration delay, we propose three

proactive resource reservation schemes, which we now detail.

#### A. Static Maximal Reservation

In this scheme, the resource reservation does not take into consideration the localization of the traffic. This corresponds to a classical scenario where higher level information (here localization) is not exploited in the lower level resource allocation. In this case, in order to preserve the reliability of URLLC, a maximal amount of resources is to be reserved, in a static, permanent manner, for URLLC slice in all the cells of the network. The optimal amount of allocated resources required for a given number of User Equipments (UEs) per cell is determined by an offline simulation which we detail in section IV-B. It is obvious that this static maximal reservation scheme will have a negative effect on eMBB throughput.

#### B. Proactive reservation on neighboring cells

In this scheme, the resource reservation for URLLC users is dynamic and proactive. Without any prior knowledge on the users' trajectory, we suppose that when an URLLC user arrives in a cell, he can move to any of the neighboring cells. Hence, a corresponding resource reservation is performed on all the neighboring cells so that, wherever the URLLC vehicle moves, his QoS is guaranteed. This is achieved as follows:

- When an URLLC UE arrives into the network, the number  $N$  of URLLC users is increased for the source and the neighboring cells, along with corresponding resource reservation obtained from the offline study.
- When the user moves from a source cell to a target cell,  $N$  is increased for the neighboring cells of the target cell that are not neighbours of the source cell, and decreased for the neighbors of the source cells that are not neighbours of the target cells, and so follows the resource reservation from the offline study.

#### C. Proactive reservation on predicted vehicular URLLC UEs trajectory

In this scenario, we suppose that the URLLC UEs' trajectory can be predicted. We can deduce for each cell the expected total number of URLLC UEs and determine the corresponding resource reservation based on the offline study. This procedure is described in Algorithm 1 where we denote by  $S$  the source cell,  $T$  the target cell, and  $X$  the destination cell that follows cell  $T$ .

This approach helps us prevent useless reservation of resources and diminishes the impact on eMBB user performance.

### IV. PERFORMANCE EVALUATION

We consider a network which consists of 13 cells forming a three-sector deployment with 500 meters inter-site distance, in compliance with the third Generation Partnership Project (3GPP) urban marco deployment (see [13]).

---

#### Algorithm 1 Trajectory dependent reservation

---

```

Create User  $i$  in cell  $S$ 
Increase by 1 in next destination  $T$ 
while User life cycle not equal to 0 do
  if Handover happens then
    Check next destination  $X$  to  $T$ 
    Increase  $N$  by 1 UE in  $X$ 
    Decrease  $N$  by 1 UE in  $S$ 
  end if
end while

```

---

#### A. Simulator description

We developed a network simulator implementing network slicing. For each cell, there are 2 slices added by default: URLLC and eMBB. For each slice, it is possible to reserve an amount of resources (PRBs) in order to achieve the target SLA requirements. The slice is created with the following properties: Slice Service Type (SST) ([14]), label, number of connected users, radio resource percentage, maximum delay and average throughput. Figure 3 illustrates the network the simulator created, showing eMBB and URLLC UEs as well as URLLC vehicle trajectory.

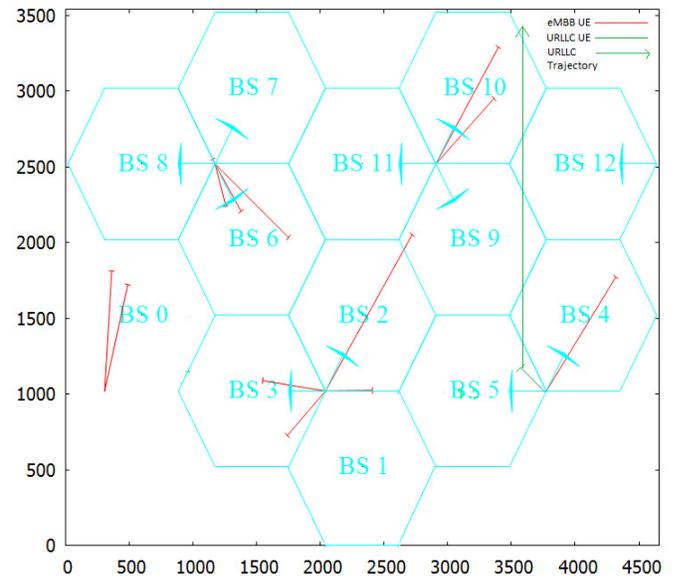


Fig. 3. Urban network with 13 gNodeBs

1) *Traffic description:* We assume that users arrive in the network following a spatial Poisson process of mean 1 [user/sec/cell] for URLLC and 3.42 [user/sec/cell] for eMBB. When users arrive, they are automatically attached to their corresponding slice. For each URLLC user, small packets of 96 bits are generated following a Poisson process with a mean 1.5 [packet/user/sec]. For eMBB, we consider a File Transfer Protocol (FTP) like traffic of fixed file size 14 Mbits. Once the file is transmitted, the eMBB user leaves the network. As for the URLLC users, they remain active for 2 minutes in average.

The eMBB users are static while the URLLC users follow a straight path with a speed of 30 km/h.

2) *URLLC Scheduling*: Both slices are multiplexed using orthogonal frequency division multiple access (OFDMA). In URLLC case, the TTI of the PRB is considered as a short one with 2 OFDM symbols or Resource Element (RE) of a duration of 0.143 msec. For eMBB, the TTI is equal to 1 ms. A PRB is extended over 14 symbols and 12 sub-carriers. Each RB has  $11 * 12 \cdot 6 = 126$  REs. All sub-carriers can be used apart from the control channels which leaves us with 126 REs.

The amount of RE needed for each URLLC UE is calculated using a look up table (LUT) obtained from system simulations. This LUT gives the coding rate using the SNR and the Block Error Rate (BLER) already calculated. Depending on the modulation order, we get:

$$bits\_per\_RE = \log_2 MO * code\_rate \quad (1)$$

where  $bits\_per\_RE$  is the number of bits applied to an RE for transmission,  $MO$  is the modulation order depending on the Modulation Code Scheme (MCS) and the code rate is the ratio between the packet size and the block size (with redundancy bits). The number of required REs is:

$$RE = file\_size / bits\_per\_RE \quad (2)$$

After calculating the needed RE, we apply the scheduling on each slice independently. During the scheduling, we calculate the total packet latency in a transmission  $L_{pack}$  based on the following equation:

$$L_{pack} = d_{tr} + d_{rt} + d_{bs} + d_{ue} + d_{al} \quad (3)$$

where  $d_{tr}$  and  $d_{rt}$  are the transmission and re-transmission times if needed, respectively, and  $d_{bs}$  and  $d_{ue}$  are the processing times of the eNodeB and UE, respectively, which, according to [15], are equal to 2.75 and 4.5 OFDM symbols, respectively.  $d_{al} = 1TTI$  is the packet alignment time. Without loss of generality, we limit the transmission to one TTI without queuing as we consider a URLLC service with very tight latency constraint and account for packet alignment and processing times as stated before.

### B. Offline optimization of the resource reservation

In order to determine the required amount of resources that have to be reserved for the URLLC slice, we perform offline simulations with a fixed amount of (static) URLLC users in each cell. The amount of resources reserved for the URLLC slice in each cell is then gradually increased until reaching the target QoS (the proportion of packets whose delay exceeds 1 ms is equal to  $10^{-6}$ ).

The obtained results indicate for instance that for 9 active users in the cell, 55% of the cell resources have to be reserved for URLLC. This level increases to 67% for 15 users.

### C. Simulation results

In this section, we compare the performances of the following scenarios in terms of URLLC reliability and eMBB throughput.

- *Allocation with 80ms reconfiguration*, described in Section II-B. We consider here a reconfiguration delay of 80 ms. This scenario is considered as a baseline as it does not implement any proactive allocation.
- *Static Maximal Reservation* corresponds to the static maximal reservation scheme described in section III-A.
- *Reservation on neighbors*, corresponds to the proactive reservation on neighboring cells described in section III-B.
- *Trajectory Prediction*, corresponds to the proactive reservation on vehicular URLLC UEs predicted trajectory described in section III-C.

In Figure 4 we illustrate the packet loss for each simulation after reaching steady state. When comparing the reservation with 80ms reconfiguration to the rest, we observe very high value of packet loss due to reconfiguration delay, while the other three scenarios, attain the requested packet loss on the order of  $10^{-6}$ .

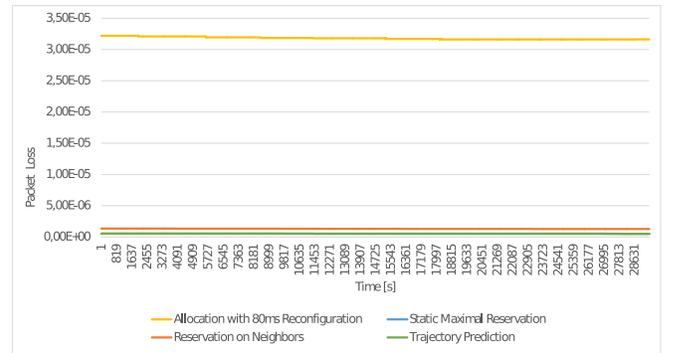


Fig. 4. URLLC packet loss

This reliability performance has an impact on eMBB throughput. As a result of over reservation in the static and neighbors reservation scenario, we can see in Figure 5 an expected degraded eMBB performance compared to the baseline where resources are reserved only when and where needed. This is due to unnecessary reservation in cells where there is a lower number or no URLLC users. Comparing the three proactive schemes, a static reservation has the largest impact on eMBB, as the maximal amount of resources is reserved for URLLC in all cells all the time. When the reservation is anticipated only in the neighbours, the impact is reduced. Finally, for trajectory prediction, the eMBB throughput is almost equal to the baseline, but with an ensured URLLC reliability.

### V. CONCLUSION AND FUTURE WORK

In this paper, we studied 5G network slicing for vehicular URLLC services and focused on the impact of slice reconfiguration delay on the performance. Having observed

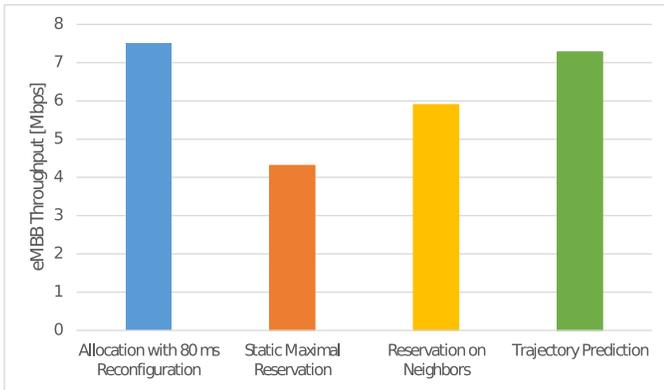


Fig. 5. eMBB average throughput for all cells in four scenarios

a URLLC QoS degradation due to this slice reconfiguration delay, we proposed proactive resource reservation schemes that anticipate slice needs for ensuring URLLC requirements. Three flavors of anticipation are considered, depending on the level of exploitation of localization information in the resource allocation procedure: a static generalized one, a scheme that applies to neighbouring cells, and one that makes use of URLLC user trajectory.

We studied the impact of these proactive schemes on eMBB performance in a scenario where URLLC and eMBB slices are sharing the same infrastructure and spectrum. We showed that with prior knowledge on the trajectory of URLLC vehicular UEs, we can limit anticipatory resource reservation and fulfill the URLLC requirements with limited impact on eMBB performance.

As future research, we aim at extending the joint management of URLLC and eMBB slices to other slice management procedures, including differentiated mobility management and slice aware load management. We also aim at implementing online learning to find dynamically the optimal resource allocation at a cell basis.

## VI. ACKNOWLEDGMENT

This research work has been funded by the EU Horizon 2020 Program under Grant Agreement No.871249 LOCUS (LOCALization and analytics on-demand embedded in the 5G ecosystem, for Ubiquitous vertical applicationS) project.

## REFERENCES

- [1] *5G White Paper*, NGMN, 2 2015, version 1.0.
- [2] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834–1853, 2018.
- [3] S. K. Sharma and X. Wang, "Toward massive machine type communications in ultra-dense cellular iot networks: Current issues and machine learning-assisted solutions," *IEEE Communications Surveys Tutorials*, vol. 22, no. 1, pp. 426–471, 2020.
- [4] *Physical layer procedures*, 3GPP, 6 2018, version 15.2.0 Release 15.
- [5] A. Anand, G. De Veciana, and S. Shakkottai, "Joint scheduling of urllc and embb traffic in 5g wireless networks," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 1970–1978.
- [6] H. D. R. Albonda and J. Pérez-Romero, "An efficient ran slicing strategy for a heterogeneous network with embb and v2x services," *IEEE Access*, vol. 7, pp. 44 771–44 782, 2019.

- [7] Y. Chen, L. Cheng, and L. Wang, "Prioritized resource reservation for reducing random access delay in 5g urllc," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–5.
- [8] K. I. Pedersen, G. Pocovi, J. Steiner, and S. R. Khosravirad, "Punctured scheduling for critical low latency data on a shared channel with mobile broadband," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–6.
- [9] *NR; Radio Resource Control (RRC) protocol specification*, 3GPP, 9 2018, version 15.3.0 Release 15.
- [10] X. Lin, D. Yu, and H. Wiemann, "A primer on bandwidth parts in 5g new radio," 2020.
- [11] T. Soni, A. R. Ali, K. Ganesan, and M. Schellmann, "Adaptive numerology—a solution to address the demanding qos in 5g-v2x," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2018, pp. 1–6.
- [12] *Study on Scenarios and Requirements for Next Generation Access Technologies*, 3GPP, 7 2020, version 16.0.0 Release 16.
- [13] *Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects*, 3GPP, 1 2010, version 1.6.0 Release 9.
- [14] *System Architecture for the 5G System*, 3GPP, 12 2017, version 15.0.0 Release 15.
- [15] *IMT-2020 self-evaluation: UP latency analysis for FDD and dynamic TDD with UE processing capability 2 (URLLC)*, 3GPP, 8 2018.