3GPP TSG RAN WG1 Meeting #AH1 **R1-1900353**
Taipei, Taiwan, January 21-25, 2019

Agenda Item: **7.2.6.3**
Source: **TCL Communication**
Title: **URLLC Configured Grant with less than K Repetitions**
Document for: **Discussion and decision**

1 Introduction

In NR, a gNB will be able to schedule pre-configured resources for URLLC users. NR supports CG transmission repetition by configuring the UE with the higher layer parameters repK and repK-RV, where repK defines the numbers of repetitions "K"of the transport block (TB) for the CG transmission and repK-RV defines the sequence of redundancy version (RV) to be used for these transmissions.

Depending on the time of arrival of the data in the buffer in relation to the periodicity P, the number of repetitions may be smaller than the configured number of repetitions K as the repeated transmissions need to stop at latest at the last transmission occasion of the period P. This scenario is shown in Figure 1. This example is for a setting where the period P comprises 4 CG occasions and the user has been configured with K (or rep-K) = 4 repetitions. The user is further configured with 4 HARQ processes where HARQ processes are determined by the resource implicitly and not communicated explicitly in the uplink control information (UCI).

In this figure, the first packet arrives at the user from the higher layers ahead of the 4 CG occasions of the current period. Thus, the user sends the 4 configured repetitions for the first packet. The 2nd packet arrives such that there are 3 occasions left in its period P and hence depending upon the RV sequence, the user can make up to 3 repetitions. The number of actual repetitions may reduce further for certain RV sequences as the starting occasions must currently start with RV 0. Further later, the third packet arrives such that only one occasion in the current period is available. This means that if the user has been configured with the suitable RV sequence, it can only make a single transmission for the current transport block. When the base station receives this single transmission for this packet, in case of decoding error, it can send a re-transmission request in the form of UL grant with the same HARQ identification (ID) for the process and new data information (NDI) set to zero, and the user will do the re-transmission for the same transport block on the assigned resources.

It would be important to highlight that the base station will be able to configure the re-transmission only if it is able to identify the user from its demodulation refence symbol (DMRS) sequence. Failing to identify the user will result in inability of base station to respond anything and the user will consider packet successfully decoded after a configured HARQ time.

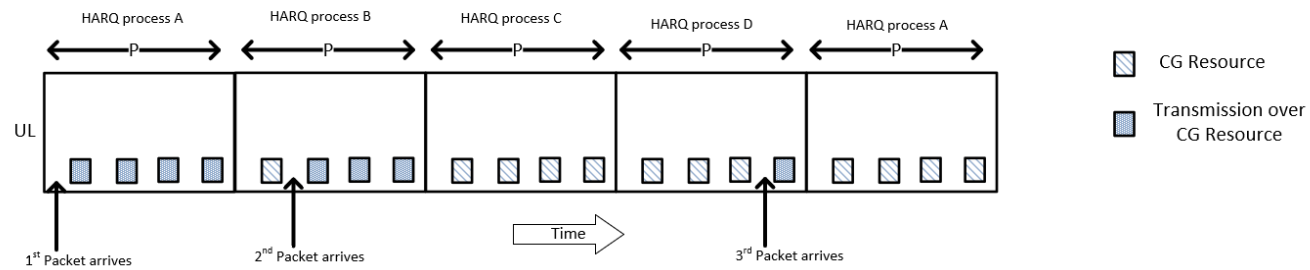

Figure 1: Problem of Less than K Repetitions

From the above examples, it becomes clear that in many cases when the packet arrival is not aligned to the appropriate timings, the user will be transmitting smaller than K repetitions on the CG occasions even if it was configured to transmit K repetitions. This imposes serious limitations on the reliability of these packets which are not transmitted for the configured number of repetitions. In many cases, this would eventually result in URLLC transmissions for which the key metric of "reliability within a certain latency" target is not satisfied.

Previous 3GPP Agreements

3GPP RAN1#95 Agreements:

- Multiple active configured grant configurations for a given BWP of a serving cell should be supported at least for different services/traffic types and/or for enhancing reliability and reducing latency
    - FFS details
    - Note: it is understood that the above may be related to RAN2-led work on intra-UE multiplexing

Although 3GPP has made this agreement but from our perspective the main motivation for multiple configurations is to support multiple services or data sizes from the UE. Thus, it's not clear whether the control and resource overhead would be acceptable to configure multiple CG configurations to handle the issue of K repetitions.

In the next section, we propose a different scheme to ensure the reliability of URLLC transmissions during the occasions when the users are able to transmit less than K repetitions.

2 Ensuring URLLC Targets for the case of less than K Repetitions

The baseline strategy to have the configured number of repetitions would be to wait for the next CG period, and transmit K repetitions. Unfortunately, this may not be the best thing to do for URLLC type of users with very strict latency and reliability targets.

To handle this performance issue with configured grant transmissions with less than K repetitions, this document proposes that when the user transmits less than K repetitions for a configured grant transmission, the transmission is supposed to require explicit HARQ feedback from the base station.

Proposal 1:

When a user is only able to transmit less than the configured number of repetitions for a configured grant transmission, this transport block will use explicit HARQ feedback structure.

The transport block sent by the user may result in three different scenarios, one for correct decoding, and two scenarios for incorrect decoding. In one of the two scenarios for incorrect decoding, the base station is able to identify the user and in the other it fails to even identify the user. We analyze in the following these scenarios and what happens in each scenario with explicit HARQ feedback structure.

1. Correct data decoding:

The receiver at the base station tries to combine all the repetitions of a transport block to facilitate data decoding, and the number of these repetitions can be less than K as per the previous discussion. Thus,

for the normal operation, the base station receiver is capable of identifying the repetitions concerning a specific TB. Thus, whenever the base station is able to correctly decode a TB, and it sees that it was sent with less than K repetitions, it will send an explicit ACK for this TB to the transmitting user.

2. Failed Data Decoding with successful UE Identification:

When the user transmits less than K repetitions, it's possible that the data decoding is not successful but the base station is able to at least identify the user transmitting the TB with less than K repetitions. This can be possible, for example, through identification of user specific DMRS sequence with which it was configured as part of CG configuration. In this case, the base station will reschedule the user for the re-transmission of the previously transmitted TB. This is the typical operation upon decoding failure with successfully UE identification, independent of if the HARQ structure is timer based or explicit feedback based.

3. Failed UE Identification:

The bad quality of received data may lead to a situation when the base station is unable to identify the transmitting user. This situation is the most damaging for the URLLC users/applications due to their tight constraints on latency and reliability. With a timer based HARQ structure, which is currently used for URLLC transmissions in 3GPP Release 15, this situation leads to different understanding at the base station and at the user. The base station, being unable to identify the user, cannot schedule the re-transmission. The user, upon receiving no UL grant for re-transmission, considers the packet successfully decoded at the base station and discards the buffer upon the expiry of HARQ timer.

Although the situation when the base station is not able to identify the user may be caused by a number of reasons, the very bad channel conditions, large amount of interference or insufficient number of actual repetitions to name a few, the configuration parameters of CG transmission, in particular MCS and the number of repetitions K, are designed so as to combat most of these adverse effects. On the other hand, if the configured number of repetitions cannot be made, this brings the CG operation point to a lower QoS target than the desired operating point.

In the proposed technique, whenever the user transmits less than K repetitions, the transport block in question is supposed to operate with explicit HARQ based feedback. In general, the base station can identify transmissions with less than K repetitions thanks to DMRS detection and CG window boundary knowledge. When the base station fails to identify the transmitting user and sends no ACK or UL grant to this user, the user upon expiry of configured HARQ timer re-transmits the TB.

Proposal 2:

For a TB, if the user initially transmits less than K repetitions on UL CG transmission mode, if it receives no ACK or UL grant for this TB within a configured HARQ timer, the user re-transmits this transport block.

The re-transmission timing and resources can be configured as part of the explicit HARQ feedback configuration. One suitable option is to re-transmit in the closest CG periodic window after the expiry of HARQ feedback timer. The HARQ feedback timer should include the time for the base station to decode the data and find the suitable occasion for potential DL transmission of HARQ ACK or UL grant for re-transmission.

In another variation of this scheme, reserved periodic resources can be made available by the base station to multiple users where they can re-transmit data in case of absence of HARQ feedback. This basically means the creation of a re-transmission zone where a group of users can be configured to re-transmit with suitable parameters.

A different scheme to combat the problem of less-than-K repetitions can be formulated. In this scheme to improve the reliability of uplink CG transmissions, whenever the user transmits less than the configured number of repetitions for a transport block, it sends the scheduling request (SR) to the base station, in parallel to transmission of transport block with less than K repetitions. This scheduling request provides a sort of diversity mechanism in parallel to the transmission of the TB.

Proposal 3:

The user is configured to transmit a scheduling request in parallel to the transmission of a TB when it is able to transmit less than K repetitions for this TB.

3 Conclusions

The followings proposals have been made in this document.

Proposal 1: When a user is only able to transmit less than the configured number of repetitions for a configured grant transmission, this transport block will use explicit HARQ feedback structure.

Proposal 2: For a TB, if the user initially transmits less than K repetitions on UL CG transmission mode, if it receives no ACK or UL grant for this TB within a configured HARQ timer, the user re-transmits this transport block.

Proposal 3: The user is configured to transmit a scheduling request in parallel to the transmission of a TB when it is able to transmit less than K repetitions for this TB.

References

[1] RAN1 Chairman's Notes, 3GPP TSG RAN WG1 Meeting #95.

[2] RAN1 Chairman's Notes, 3GPP TSG RAN WG1 Meeting #94Bis.

[3] RAN1 Chairman's Notes, 3GPP TSG RAN WG1 Meeting #94.