**t3GPP TSG-RAN WG2 Meeting #114 R2-210xxxx**

**Online, May 19 – May 27, 2021**

**Source: CATT**

**Title: Summary of Email Discussion 506 – R17 IIOT QoS**

**Agenda Item: 8.5.4**

**Document for: Discussion and Decision**

# Introduction

This contribution provides a summary of the following email discussion:

* [POST113bis-e][506][R17 IIoT] Enhancements based on QoS (CATT)

Scope:

* + - Discuss remaining open issues related to RAN enhancements on new QoS based on inputs submitted to AI 8.5.4.
		- Agreeable Proposals

 Deadline: same as submission deadline (as set per RAN2 chair)

Rapporteur would like to set the deadline for companies inputs to 05-10-2021 23:00 UTC to have enough time for preparing the summary report.

# Contact information

|  |  |
| --- | --- |
| Company | Name and email address |
| CATT | Pierre Bertrand; pierrebertrand@catt.cn |
| Ericsson | Zhenhua Zou; zhenhua.zou@ericsson.com |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Discussion

As a follow-up of the offline #506 [2], the following agreements on RAN enhancements based on new QoS were achieved in RAN2#113-e [1]:

**Agreements**

- Communication service availability (CSA) is not needed on top of survival time. Send a reply LS to SA2 to notify such confirmation

*-* RAN2 confirms that specification enhancement for survival time support may only needed for uplink. Downlink is addressed by implementation and no specification impacts.

*-* Support for survival time in UCE is up to network configuration.

- Communication service reliability (CSR) is not needed on top of survival time

- Only periodic traffic is considered for survival time work in Rel-17

- RAN2 assumes one application message is conveyed by one PDCP SDU, and may further consider the cases where one application message is conveyed by varying number of PDCP SDUs depending on the progress

Considering the above agreements, and based on the contributions posted at RAN2#113-e [5]-[22], we suggest the following plan for this email discussion:

* Survival Time
	+ Survival Time trigger
	+ Link reliability increase mechanisms
	+ Survival Time management
* Burst spread / BET

## Survival Time

As agreed in RAN2#113bis-e, the relevant Survival Time requirements to consider are those of the Periodic deterministic communication service which are expressed by SA1 in Table 5.2-1 from TS 22.104 [3], copied below for convenience/reference.

TS 22.104 [3] defines Survival Time (ST) as “the time that an application consuming a communication service may continue without an anticipated message.” Figure C.3-1 in Annex C.3 of the TS 22.104 provides an illustration of the survival time.

Moreover, SA2 finalized in SA2#143E the inclusion of the Survival Time parameter for a given TSC flow in the TSCAI in TS23.501 [4] as follows, referring to the definition in TS 22.104:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Survival Time may be provided by TSN AF/AF either in terms of maximum number of messages or in time units, where a time unit is equivalent to the Periodicity. During a single period, single burst is expected. If Survival Time is provided in terms of maximum number of messages, the SMF coverts it into time units by multiplying its value by the Periodicity provided in the TSCAI Container. The SMF corrects the Survival Time in time units by the previously received cumulative rateRatio from the UPF and sets the TSCAI Survival Time to the corrected value.Table 5.27.2-1: TSC Assistance Information (TSCAI)

|  |  |
| --- | --- |
| Assistance Information | Description |
| Flow Direction | The direction of the TSC flow (uplink or downlink). |
| Periodicity | It refers to the time period between start of two bursts. |
| Burst Arrival time (Optional) | The latest possible time when the first packet of the data burst arrives at either the ingress of the RAN (downlink flow direction) or egress interface of the UE (uplink flow direction). |
| Survival Time (Optional) | It refers to the time period an application can survive without any burst, as defined in Annex C.2.3, TS 22.104 [105]. |

 |

Table 5.2-1: Periodic deterministic communication service performance requirements

| Characteristic parameter | Influence quantity |  |
| --- | --- | --- |
| Communica­tion service availability: target value (note 1) | Communication service reliability: mean time between failures | End-to-end latency: maximum (note 2) (note 12a) | Service bit rate: user experienced data rate (note 12a) | Message size [byte] (note 12a) | Transfer interval: target value (note 12a) | Survival time (note 12a) | UE speed (note 13) | # of UEs | Service area (note 3) | Remarks |
| 99.999 % to 99.999 99 % | ~ 10 years | < transfer interval value | – | 50 | 500 μs  | 500 μs | ≤ 75 km/h | ≤ 20 | 50 m x 10 m x 10 m | Motion control (A.2.2.1) |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < transfer interval value | – | 40 | 1 ms  | 1 ms | ≤ 75 km/h | ≤ 50 | 50 m x 10 m x 10 m | Motion control (A.2.2.1) |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < transfer interval value | – | 20 | 2 ms  | 2 ms | ≤ 75 km/h | ≤ 100 | 50 m x 10 m x 10 m | Motion control (A.2.2.1) |
| 99.999 9 % | – | < 5 ms | 1 kbit/s (steady state)1.5 Mbit/s (fault case) | < 1,500 | < 60 s (steady state)≥ 1 ms (fault case) | transfer interval | stationary | 20 | 30 km x 20 km | Electrical Distribution – Dis­tributed automated switch­ing for isolation and service restoration (A.4.4); (note 5)  |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < transfer interval value |  | 1 k | ≤ 10 ms | 10 ms | - | 5 to 10 | 100 m x 30 m x 10 m | Control-to-control in motion control (A.2.2.2); (note 9) |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < transfer interval value (note 5) | 50 Mbit/s |  | ≤ 1 ms | 3 x transfer interval | stationary | 2 to 5 | 100 m x 30 m x 10 m | Wired-2-wireless 100 Mbit/s link replacement (A.2.2.4) |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < transfer interval value (note 5) | 250 Mbit/s |  | ≤ 1 ms | 3 x transfer interval | stationary | 2 to 5 | 100 m x30 m x 10 m | Wired-2-wireless 1 Gbit/s link replacement (A.2.2.4) |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < transfer interval value |  | 1 k | ≤ 50 ms | 50 ms | - | 5 to 10 | 1,000 m x 30 m x 10 m | Control-to-control in motion control (A.2.2.2); (note 9) |
| > 99.999 9 % | ~ 10 years | < transfer interval value | – | 40 to 250 | 1 ms to 50 ms (note 6) (note 7) | transfer interval value | ≤ 50 km/h | ≤ 100 | ≤ 1 km2 | Mobile robots (A.2.2.3) |
| 99.999 9 % to 99.999 999 % | ~ 1 month | < transfer interval value | – | 40 to 250 | 4 ms to 8 ms (note 7) | transfer interval value | < 8 km/h (linear movement) | TBD | 50 m x 10 m x 4 m | Mobile control panels – remote control of e.g. assembly robots, milling machines (A.2.4.1); (note 9) |
| 99.999 999 % | 1 day | < 8 ms(note 14) | 250 kbit/s | 40 to 250 | 8 ms | 16 ms | quasi-static; up to 10 km/h | 2 or more | 30 m x 30 m | Mobile Opera­tion Panel: Emer­gency stop (connectivity availability) (A.2.4.1A) |
| 99.999 99 % | 1 day | < 10 ms(note 14) | < 1 Mbit/s | <1024 | 10 ms | ~10 ms | quasi-static; up to 10 km/h | 2 or more | 30 m x 30 m | Mobile Operation Panel: Safety data stream (A.2.4.1A) |
| 99.999 999 % | 1 day | 10 ms to 100 ms(note 14) | 10 kbit/s | 10 to 100 | 10 ms to 100 ms | transfer interval | stationary | 2 or more | 100 m² to 2,000 m² | Mobile Operation Panel: Control to visualization (A.2.4.1A) |
| 99.999 999 % | 1 day | < 1 ms(note 14) | 12 Mbit/s to 16 Mbit/s | 10 to 100 | 1 ms | ~ 1 ms | stationary | 2 or more | 100 m² | Mobile Operation Panel: Motion control (A.2.4.1A) |
| 99.999 999 % | 1 day | < 2 ms (note 14) | 16 kbit/s (UL)2 Mbit/s (DL) | 50 | 2 ms | ~ 2 ms | stationary | 2 or more | 100 m² | Mobile Operation Panel: Haptic feedback data stream (A.2.4.1A) |
| 99.999 9 % to 99.999 999 % | ~ 1 year | < transfer interval | – | 40 to 250 | < 12 ms (note 7) | 12 ms | < 8 km/h (linear movement) | TBD | typically 40 m x 60 m; maximum 200 m x 300 m | Mobile control panels -remote control of e.g. mobile cranes, mobile pumps, fixed portal cranes (A.2.4.1); (note 9) |
| 99.999 9 % to 99.999 999 % | ≥ 1 year | < transfer interval value | – | 20 | ≥ 10 ms (note 8) | 0 | typically stationary | typically 10 to 20 | typically ≤ 100 m x 100 m x 50 m | Process automation – closed loop control (A.2.3.1) |
| 99.999 % | TBD | ~ 50 ms  | – | ~ 100 | ~ 50 ms | TBD | stationary | ≤ 100,000 | several km2 up to 100,000 km2 | Primary frequency control (A.4.2); (note 9) |
| 99.999 % | TBD | ~ 100 ms | – | ~ 100 | ~ 200 ms | TBD | stationary | ≤ 100,000 | several km2 up to 100,000 km2 | Distributed Voltage Control (A.4.3) (note 9) |
| > 99.999 9 % | ~ 1 year | < transfer interval value | – | 15 k to 250 k | 10 ms to 100 ms (note 7) | transfer interval value | ≤ 50 km/h | ≤ 100 | ≤ 1 km2 | Mobile robots – video-operated remote control (A.2.2.3) |
| > 99.999 9 % | ~ 1 year | < transfer interval value | – | 40 to 250 | 40 ms to 500 ms (note 7) | transfer interval value | ≤ 50 km/h | ≤ 100 | ≤ 1 km2 | Mobile robots (A.2.2.3) |
| 99.99 % | ≥ 1 week | < transfer interval value | – | 20 to 255 | 100 ms to 60 s (note 7) | ≥ 3 x transfer interval value | typically stationary | ≤ 10,000 to 100,000 | ≤ 10 km x 10 km x 50 m | Plant asset management (A.2.3.3) |
| >99.999 999 % | > 10 years | < 2 ms | 2 Mbit/s to 16 Mbit/s | 250 to 2,000 | 1 ms | transfer interval value | stationary | 1 | < 100 m2 | Robotic Aided Surgery (A.6.2) |
| >99.999 9 %  | > 1 year | < 20 ms | 2 Mbit/s to 16 Mbit/s | 250 to 2,000 | 1 ms | transfer interval value | stationary | 2 per 1,000 km2 | < 400 km (note 12) | Robotic Aided Surgery (A.6.2) |
| >99.999 % | >> 1 month (< 1 year) | < 20 ms | 2 Mbit/s to 16 Mbit/s | 80 | 1 ms | transfer interval value | stationary | 20 per 100 km2 | < 50 km (note 12) | Robotic Aided Diagnosis (A.6.3) |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < 0.5 x transfer interval | 2.5 Mbit/s | 250500 with localisa­tion informa­tion | > 5 ms> 2.5 ms> 1.7 ms(note 10) | 0transfer interval2 x transfer interval(note 10) | ≤ 6 km/h (linear movement) | 2 to 8 | 10 m x 10 m x 5 m;50  m x 5 m x 5 m(note 11) | Cooperative carrying – fragile work pieces; (ProSe communication) (A.2.2.5) |
| 99.999 9 % to 99.999 999 % | ~ 10 years | < 0.5 x transfer interval | 2.5 Mbit/s | 250500 with localisa­tion informa­tion | > 5 ms > 2.5 ms> 1.7 ms (note 10) | 0transfer interval2 x transfer interval(note 10) | ≤ 12 km/h (linear movement) | 2 to 8 | 10 m x 10 m x 5 m;50 m x 5 m x 5 m(note 11) | Cooperative carrying – elastic work pieces; (ProSe communication) (A.2.2.5) |
| NOTE 1: One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement.NOTE 2: Unless otherwise specified, all communication includes 1 wireless link (UE to network node or network node to UE) rather than two wireless links (UE to UE).NOTE 3: Length x width (x height).NOTE 4: (void)NOTE 5: Communication includes two wireless links (UE to UE).NOTE 6: This covers different transfer intervals for different similar use cases with target values of 1 ms, 1 ms to 10 ms, and 10 ms to 50 ms.NOTE 7: The transfer interval deviates around its target value by < ±25 %.NOTE 8: The transfer interval deviates around its target value by < ±5 %.NOTE 9: Communication may include two wireless links (UE to UE).NOTE 10: The first value is the application requirement, the other values are the requirement with multiple transmission of the same information (two or three times, respectively).NOTE 11: Service Area for direct communication between UEs. The group of UEs with direct communication might move throughout the whole factory site (up to several km²). NOTE 12: Maximum straight-line distance between UEs. NOTE 12a: It applies to both UL and DL unless stated otherwise.NOTE 13: It applies to both linear movement and rotation unless stated otherwise. NOTE 14: The mobile operation panel is connected wirelessly to the 5G system. If the mobile robot/production line is also connected wirelessly to the 5G system, the communication includes two wireless links. |

Therefore, from RAN perspective, it should be a common understanding that, when a service flow, configured with Survival Time, enters Survival Time, RAN action should be to improve the associated link reliability to make sure any subsequent message(s) can be delivered successfully before the Survival Time is violated. It is therefore proposed to confirm this baseline understanding.

**Q1: Do you agree with the basic principle that when a service flow, configured with survival time, enters survival time, RAN action (gNB and/or UE) should be to improve the associated link reliability to make sure any subsequent message(s) can be delivered successfully before the survival time is violated?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | Yes, with comments | Survival time information is optional, and it is not required that the network must use it, e.g., monitor the survival time and “improve” the link quality subsequently. But we agree that it is how it can be used. One additional comment is that, once survival time mode is entered we cannot “make sure” the message is delivered before survival time expiry, we can only allow the option of “improving the associated link reliability” to meet its service level requirement, e.g., communication service availability, communication service reliability and etc. A more precise wording can be that: When a survival time information is provided in TSC AI, RAN action (gNB and/or UE) can utilize it to monitor survival time and, when entering survival time, improve the associated link reliability so that any subsequent message(s) can be delivered to meet its service level requirement.  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

### Survival Time Trigger

From the contributions posted at RAN2#113bis-e, the proposed solutions for triggering Survival Time “state” can be classified into three main categories:

1. Left to gNB implementation only, so that specification enhancements for UE-based solutions are not needed: [14][17][19][20]
2. UE-based with proactive trigger: UE triggers ST based on Sequence Number i.e. boosts reliability proactively every N packets [5][7]
3. UE-based with reactive trigger: UE triggers ST based on Tx-side timer or Tx-side counting of transmission failures, and autonomously applies a more reliable pre-configured UL transmission [5][6][8][9][11][12][15][16][17][18][21][22]

We will address each category in the following.

#### Survival Time handling left to gNB implementation only

Clearly this question needs to be sorted out first, because if ST handling can be fully addressed by gNB implementation, no further work is expected in RAN2 for this topic.

This proposal essentially relies on the observation that, even for the most stringent TSN flows of above Table 5.2-1 (ST = 0.5ms), [14] “*for UL traffic, the reception status is known at the gNB first and so there is no faster and more reliable way to indicate to the UL transmission reception status to the UE than gNB simply dynamically scheduling the UE*”. “*Upon observing this a gNB can schedule the subsequent packet with higher reliability to help ensure the survival time is not violated, such as, sending a (re)-activation command for UL CG or a dynamic uplink grant with a more robust MCS, or even activating PDCP duplication.*”

On the other hand, [6] specifically assesses this solution with a quantitative analysis of the reaction time for the most stringent usecase of Table 5.2-1 (ST = 0.5ms), which confirms that, as proposed in [14], gNB has enough time to provide either a (re)-activation command for UL CG or a dynamic uplink grant with a more robust MCS, but still, [6] lists the following drawbacks associated with such a solution for such usecase:

* this only leaves CG type 2 reconfiguration as solution for improving the reliability and rules out other solutions e.g. duplication activation via MAC CE
* CG type 2 reconfiguration can only play with MCS for improving the reliability, but that may not be sufficient to address deep link quality decrease due to e.g. beam blockage (which is why duplication was designed for NR in first place)
* it can only work with Type 2 CGs, not Type 1 CGs
* in this case gNB must schedule the retransmission grant *after* reconfiguring the CG, since the two commands are addressed to CS-RNTI with NDI = 1 and 0 respectively, hence must be serialized.
* this requires configuring at least 3 HARQ processes for this CG configuration, although only 2 are required in legacy implementation (see [6] Annex 2).

In addition, [15] brings the following concern with leaving ST handling all to gNB implementation: “*We feel that it is essential that a basic framework for ST handling is standardized. While entering the ST state is not expected to be a frequent event, when it does happen it is absolutely crucial that the procedure is executed in a standardized, verifiable manner due to requirements of applications for which it will most often be configured*”.

In summary, it is obvious that gNB has full freedom to use legacy scheduling and reconfiguration mechanisms to upgrade the radio resources targeted for an UL flow whenever it estimates it is needed. Therefore, if it can be considered that gNB implementation can be sufficient to address *some* of the usecases in the above Table 5.2-1, the key point to clarify is whether it is sufficient to handle all usecases of Table 5.2-1.

**Q2: Do you agree that handling of Survival Time in RAN can be all left to gNB implementation only so that specification enhancements for UE-based solutions are not needed?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | Yes | Comments to the drawbacks listed in the paper [14][6]1. There is no requirement whatsoever to deliver the already “failed” message that triggering the survival time mode, and what it matters is the subsequent message. In other words, there is no need for the gNB to schedule two serialized DCI commands or use three HARQ processes.
2. We acknowledge that network implementation cannot work with CG type 1, but CG type 1 was introduced in Rel-15 to reduce the latency when the network initially configures and activates the Configured Grant. If the latency in activation is an issue, network would obviously use CG type 2 in this case. In other words, it is quite strange to enhance a feature that was not intended for this use case and for which there is a clear better alternative (CG type 2).
3. There is a network implementation solution to activate PDCP duplication for survival time by sending a DCI command. Network can configure and activate PDCP duplication in a second cell, and the duplicate RLC entity on this cell is restricted by LCP restriction configuration to be sent on a configured but de-activated CG. The PDCP discard timer is set equal to the PDB. The network can send a DCI activation command to activate this CG if it sees the need. Since the PDCP discard timer is set equal to PDB, when the CG is activated, only the latest PDCP duplicate packet from the UE will be transmitted.
4. For the comment in [15], the standardization is to provide essential features for inter-operability between different UE and network vendors. It is not to standardize the network behaviour.

The bottom line here is, for medium and long survival time, it can be up-to gNB implementation and what is contentious is the case for short survival time (e.g., 0.5 millisecond). In this case, our argument is, for short survival time, the UE-based reactive triggering is no better than gNB implementation, since the feedback is given by gNB. More importantly, we see there are some fundamental technical flaws in all those UE-based reactive triggering solutions, see more inputs below. For the UE-based proactive triggering, it can also be implemented by a network implementation solution, see details below. Last but not the least, it is not required that the network must use the survival time mechanism to deal with this very stringent survival time. gNB is the first to know when an UL transmission expected within the context of a CG has failed and should be responsible for starting the ST timer, determining what to do (if anything) after starting ST timer and what to do (if anything) when ST timer expires. Any use case for which short periodicity precludes the ability of a gNB to dynamically react to detecting an UL transmission failure (by allocating additional UL resources for UE re-transmission before ST expiry) can also be left up to gNB implementation i.e. a gNB should be able to determine that there are CG type applications for which satisfying PER alone is sufficient (i.e. the gNB does nothing upon detecting ST expiry).  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### UE-based with proactive trigger

As described in [7], in this approach, UE “*does not rely on any kind of feedback (e.g. new dynamic grant, re-transmission grant, or ARQ NACK) from the receiver (gNB), but the UE proactively boost the reliability of at least one burst in every N-th incoming burst to make sure consecutive error of N burst does not occur. The PDCP layer may directly determine how to deal with a incoming packet based on its sequence number (SN)*”.

Specifically, [7] elaborates on the benefits of such approach compared with *reactive* triggers for which associate issues are listed in observations 1-3 as follows:

* Relying on feedback (e.g. HARQ feedback or uplink grant) for survival time support would end up a very restrictive solution which basically precludes possibility of flexible TDD, and any small jitter in the end-to-end path would make this solution useless due to extremely tight time budget.
* If PDCP duplication is to be used as the mean to boost reliability, relying on feedback (e.g. HARQ feedback or uplink grant) for survival time support would result in complicated cross-layer interaction that breaks the current way of RAN protocol operation, as PDCP needs to check status of HARQ processes.
* Relying on feedback (e.g. HARQ feedback or uplink grant) for survival time support may require RAN3 involvement as information exchange between UPF/SMF and gNB may be needed to enable cross link coordination for end-to-end survival time support in UE-to-UE communications.

And, regarding the lack of efficiency due to the resulting resource waste brought up in previous discussions as a drawback of this method, [7] brings the following argument:

*“One may argue that Option 4 is not efficient because reliability of some packets may be boosted unnecessarily. Nevertheless, for schemes based on feedback and autonomous PDCP duplication, the gNB would anyway have to pre-allocate configured grant (CG) resources for the duplication leg, in order to ensure that duplication leg can have radio resource immediately after it is activated. Hence, this is not more spectrally efficient as compared to proactive schemes. Moreover, it is worth noting that in an IIoT/TSC use case, it is much more important to ensure applications do not fail frequently than optimizing spectral efficiency, because failure of applications in such use cases can be quite costly. Thus, spectral efficiency should not be seen as the most prominent criteria when considering the mechanism for supporting survival time.”*

**Q3: Do you support Survival Time triggered proactively based on Sequence Number?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | In addition to that there is always a resource waste for the “proactively” allocated resources (only one out of 10^4 to 10^5 would be actually useful if one assumes PER of 10^-4 and 10^-5), this can be done by gNB implementation. For example, the network can configure a duplicated leg restricted on a CG with periodicity=N\*message periodicity and PDCP discard timer is set to PDB.  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### UE-based with reactive trigger

As mentioned in Section 3.1.1, this approach mainly includes two alternate options, either based on Tx-side timer or Tx-side counting of transmission failures that we elaborate further down. In short, with reactive triggers, UE autonomously applies a more reliable pre-configured UL transmission only when needed, i.e. when it *actually* experiences Survival Time, and according to the detailed latency analysis in [6], it meets the requirements for the most stringent TSN flows of TS22.104 reusing the legacy CG configuration for that flow (no need for additional HARQ process to be configured) and allows a wide range of transmission reliability increase methods (duplication, L1/L2 parameters adaptation, …) based on pre-configuration. Note it is Rapporteur’s understanding that the additional pre-configured resources potentially used during Survival Time (e.g. with PDCP Duplication) are not wasted in “normal” state if NW knows when they are used/not used by the UE i.e. if NW knows when UE triggers Survival Time, (e.g. based on NW feedback) in which case it can reuse them in “normal” state.

**Q4: Do you support UE-based reactive ST trigger?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson  | No  | We are not sure if we understand the comment by the rapporteur that the additional pre-configured resources are not wasted. If there is no other traffic to be transmitted from the UE, then the resources are wasted. In addition, the allocated resources are configured for URLLC traffic and, thus, spectrum inefficient for eMBB traffic. Similar to the response to Q3, this would always result in an unacceptable waste of UL resources, since the network has to pre-allocate all these physical resources so that the UE can use it without network indication in, e.g., one out of 10^4 to 10^5 cases. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### Tx-side Timer

This solution is supported in contributions [8][10][11][12][15][17][18].

The triggering timer, e.g. equal to 5G-AN PDB [11][12][17], is (re)started either at the reception of the packet in PDCP [11][12] or at BAT [17], or upon receiving ACK for the previous packet [15][18] and stopped upon receiving ACK. Survival Time is triggered when such timer expires [11][12][15][17][18] and optionally also upon receiving NACK [11][12][17]. It requires gNB always sending either ACK or NACK for each packet. It ensures that a transmitter interprets a lack of feedback (PDCCH miss) as a possible failed transmission but may lead to uselessly adapting the transmission resource configuration to achieve higher reliability (e.g. in case missed PDCCH was carrying ACK).

**Q4-a: If you answered “Yes” to Q4, do you support Tx-side Timer method for triggering ST?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | Even though we haven’t answered “Yes”, we would like to point out the issues in this solution.1. This solution requires gNB to send always a ACK and NACK, which is not feasible in licensed spectrum operation, as there is no PHICH channel in NR. For shared spectrum operation, it is agreed in the last meeting that the support in UCE is up-to network configuration.
2. As strongly argued by the rapporteur that gNB based implementation has issues when the survival time is short, the tx-side timer has similar issues since transmitting a PDCCH every 0.5 millisecond is a large resource waste, if feasible at all.
 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### Tx-side Counter

This solution is supported in contributions [5][6][8][16][18][21][22].

Survival Time is triggered when UE experiences N consecutive UL transmission failures for the flow/DRB/LCH configured with Survival Time. Transmission failures are detected by receiving NACK from the NW. NACK can be HARQ-NACK, e.g. DCI retransmission grant [5][6][8][16][18], or RLC NACK [5][16]. In its simplest form, N=1 and Survival Time is triggered upon receiving a HARQ NACK [5][6][8][16][18].

**Q4-b: If you answered “Yes” to Q4, do you support Tx-side counter-based method for triggering Survival Time?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | Even though we haven’t answered “Yes”, we would like to point out the issues in this solution.1. The network may dimension multiple retransmission grants to meet the (PDB, PER) target in the “normal” state. It is not clear from which retransmission grant the UE needs to re-act.
2. This method mandates that the network to always schedule a retransmission for the “failed” message transmission, which is far less important than the message in the subsequent period.
3. As strongly argued by the rapporteur that gNB based implementation has issues when the survival time is short, the tx-side counter would not work with RLC NACK and a short survival time.
 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### Other triggers

#### Feedback enhancements

In [12][22], it is suggested that new feedback mechanisms, faster than current HARQ RTT should be studied. For example, in [12]: “*in order to be able to accurately determine whether the data packet has been successfully sent within the AN PDB time, the feedback corresponding to the data packet needs to be notified to the sender within the AN PDB time and as soon as possible. This part may still need to be enhanced*”. And in [22]: “*To meet the stringent delay requirement of the URLLC services, fast feedback should be introduced to reduce the HARQ feedback delay*”.

However, Rapporteur suggests, to narrow the scope of this discussion, that the above two enhancements can be considered as part the above Tx-side timer and/or counter based solutions, if selected, and can be discussed later.

#### Survival Time and UCE

In RAN2#113-e, it was agreed:

*-* Support for survival time in UCE is up to network configuration.

[10] asks RAN2 to consider if survival time can be applicable to UCE. However, it is Rapporteur’s understanding that one consequence of the above agreement it that configuring/scheduling UL transmissions of a flow/DRB/LCH configured with Survival Time on an unlicensed frequency cell under controlled environment (UCE) is not precluded and so, should be supported. Then, [5] suggests that Survival Time in UCE should be handled differently than in non-UCE, for example leveraging the *cg-RetransmissionTimer* and taking LBT failures into account. At this early stage, rather than going into the details, we propose first to answer the below basic question.

**Q5: Should we consider additional aspects for handling Survival Time in UCE?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | We are not sure why we re-discuss the same question, since the paper [10] is a resubmission whose topic was discussed and concluded in the last meeting. In our view, industrial IIoT would mostly and firstly be deployed in the licensed spectrum and we want to focus on enhancements for licensed spectrum first. At this moment, with the assumption of UCE (i.e., LBT fails rarely), the operation LBT failures can increase the risk of ST expiration they are seen as occurring infrequently and therefore we can justify not considering their potential impact at least within the scope of Rel-17 ST operation. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

### Link reliability increase mechanisms

Consistent with the outcome of the RAN2#113-e meeting offline [2], the solutions proposed in the contributions to RAN2#113bis-e can be classified into 3 groups:

1. Duplication activation [5][6][7][8][9][13][15][16][17][22]
2. L1/L2 configuration adaptation [5][7][8][10][11][15][16]
3. gNB scheduling [14][17][19][20]

However, we think the gNB scheduling option is already addressed by Q1. Indeed, similar to what we mentioned in Section 3.1.1.1, it is obvious that gNB has full freedom to use legacy scheduling and reconfiguration mechanisms to upgrade the radio resources targeted for an UL flow whenever it estimates it is needed, and the question here is rather whether it is sufficient to handle all usecases of Table 5.2-1, which is already addressed by Q1. Therefore, we focus here on the link reliability increase mechanisms autonomously triggered by UE-based solutions.

#### Duplication activation

**Q6: Should it be possible for a UE, when configured accordingly, to autonomously activate duplication when Survival Time is triggered?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | Per Q2, we propose a gNB can and should determine when the ST timer is triggered and what to do thereafter (if anything). It would be possible if network has pre-allocated the resource for UE to transmit the duplicates. But it requires to have reserved UL resources all the time for the LCHs with PDCP duplication configured but not activated yet. Use of survival time is a very rare event which means that any reserved UL resources would ONLY be used for one out of 10^4 to 10^5 attempted message transmissions, making this approach extremely spectrum inefficient. In this aspect, we don’t consider it is “possible” since it does not make sense for the network to allow this possibility.  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### L1/L2 configuration adaptation

At this stage, this option can include, for example, flexible L2 configuration switching (e.g. RLC leg switching [7][8]), dynamic LCH configuration change (e.g. priority), UE autonomous LCH restriction relaxation, and LCP adjustment by pre-defined rule [10], adaptive L1 configuration [11] e.g. lower the MCS or boost the data transmission power [16].

**Q7: Should it be possible for a UE, when configured accordingly, to autonomously adapt its L1/L2 configuration when Survival Time is triggered?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | See response to Q6. In addition, there are no details for each solution. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

### Survival Time management

Other than increasing the reliability, the UE behavior while in Survival Time is discussed in various contributions, specifically:

1. Survival Time timer
2. Return to normal state

#### Survival Time timer

A timer is discussed for monitoring the Survival Time itself [5][8][11][15][18]: the transmitter starts the Survival Time timer upon one of the triggers discussed in Section 3.1.1 (whichever is agreed). The transmitter would subsequently stop and reset the timer when indicated that one (or more) transmission(s) succeeded via e.g. explicit HARQ (or RLC [15]) ACK or upon CG timer expiry [5]. However this timer only makes sense if UE has an action upon its expiry, which is proposed to be studied by [18]. The proposed UE behaviors discussed in contributions depending on Survival Time timer status are:

* UE decreases / increases reliability if it receives ACK while timer is running / not running (incl. expired), respectively. Or the other way around. [15]
* UE returns to normal state when it leaves Survival Time (understand incl. timer expiry) [11]
* There is no specific behaviour upon Survival Time expiry, so such timer is not needed [5]

**Q8: Is a Survival Time timer needed for monitoring the Survival Time itself? If yes, please indicate the impact of this timer on the UE behavior (e.g. when timer expires).**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | Per Q2 response, the gNB is in control of what to do when ST timer is started and expires 🡪 the UE does not need a new timer. We are not sure if there is any difference from the “Tx-side Timer” discussed in the subsection 3.1.1.3.1. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

#### Return to normal state

Once a UE is in Survival Time, [8][11][17][18] suggest that it autonomously returns to normal state (default reliability) when indicated that one (or more) transmission(s) succeeded. However, it is expected that those companies that consider that Survival Time trigger should all be left to gNB implementation would also assume it is under gNB responsibility to decrease the link reliability when it is appropriate to do so. Moreover, given it might be considered less time critical as for entering Survival Time, even with a UE-based Survival Time trigger, one could consider that the relaxing of the UL transmission reliability can be left under gNB control.

**Q9: Should UE autonomously return to normal state (default reliability) based on pre-configured criterion, or should it be left to network’s responsibility when and how to reconfigure the UL transmission back to a default configuration?**

|  |  |  |
| --- | --- | --- |
| **Company** | **UE/NW** | **Comments** |
| Ericsson | NW | It is indeed less time critical to exit the survival time and can be left for network control.  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

## Burst Spread, Burst End Time

During SA2#143E, no conclusion could be reached regarding the inclusion of the Burst Spread and Burst End Time (BET) parameters in TSCAI parameters, hence they are not captured in TS 23.501 [4]. As a result, if RAN2 decides that either or both parameter is needed by RAN, RAN2 should explicitly ask SA2 to (re)consider introducing the(se) parameter(s) in TSCAI.

### Burst Spread

As a consequence of SA2 decision (or lack of), [5][7][14] suggest to not consider the Burst Spread parameter in RAN, [10][17] to further wait for SA2, while [8] think it should not impact RAN specifications anyways.

**Q10: Should RAN2 ask SA2 to consider Burst Spread in TSCAI?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | No | Not needed. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

### BET

BET is promoted in [14] as a key parameter supplementing the burst arrival time (BAT) thus allowing a gNB to determine the latest point in time where UE or gNB can receive a packet for inclusion in the next instance of a radio resource (uplink or downlink respectively) used to support transmission of packets associated with a given QoS flow. Further arguments in support of BET are given in [14] as follows:

* The key benefit of BET is that it provides the gNB with enhanced flexibility regarding where it can schedule the DRB resources needed for supporting a given QoS flow while satisfying the PDB requirement. Without BET a gNB will not know the point in time when the last packet will arrive which forces it to identify a latest point in time (i.e. a “drop dead time”) for accepting packets to be aggregated into the next MAC PDU to be sent, e.g., using a periodic DL SPS configuration or in a dynamic resource allocation used for the corresponding QoS flow. Assuming a “drop dead time” for BET can substantially reduce the flexibility regarding where a gNB can schedule the corresponding DRB resources and thereby reduce overall efficiency in managing radio resources. When there is only one packet in the burst, the value of the BET is the same as the value of BAT.
* The value of BET can be determined e.g. using the width of the TSN gate open cycle during which the UPF/NW-TT receives the set of packets and therefore enhancing TSCAI to include BET seems to be a feasible option.

In addition, [14] considers that BET is necessary for a correct Survival Time monitoring because the survival time should be counted from the PDB after the burst end time, see below Figure 1. In this way, it accounts for the possible latest arrival time in the subsequent bursts and thus the survival time can make sense for RAN.

 

Figure 1 Correct Survival Time Calculation

Opposite views are expressed in [7][8][17][20] arguing that BET can be deduced from other parameters e.g. burst size (or MDBV), while TSC flows are expected to experience minimum jitter, or that time difference between burst arrival time and burst end time is not expected to be large enough to be taken into account by the gNB for scheduling. And [5] considers that, following a reactive approach where RAN would start the survival state **reactively** upon a packet failure without considering burstiness of the application (Tx-side counter, Section 3.1.1.3.2) Burst Ending Time and Burst Spread knowledge are not needed at RAN for survival time monitoring.

**Q11: Should RAN2 ask SA2 to consider BET in TSCAI?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| Ericsson | Yes | It is our understanding that MDBV (which is used for burst size in TSC AI) refers to the **maximum** data burst volume. If the actual message size is smaller than MDBV, then it is infeasible to deduce BET. Only if the actual message size is always equal the MDBV, then the network can deduce the BET. But the whole point of the TSC AI with BAT is for the network to pre-allocate the resources before traffic arrival. Otherwise, network can observe the arrival time and there is no need for BAT either. We cannot assume that the time difference between burst arrival time and burst end time will never be large enough to be taken into account by the gNB for scheduling. In addition, if a gNB always assumes BET = BAT, then the gNB may allocate DRB resources that start prior to arrival of the last UL/DL packet to be included in a given CG/SPS resource and there are no resources for the transmission of these packets.Regardless of the outcome of this discussion, we believe that there is a need to clarify that when the survival time starts, for example, if the BET were not agreed in this release, then at least the start of the survival time should be at some definite timing point, such as BAT + PDB. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

## Other

The scope of this email discussion is intended to be kept rather reduced due to the small time period until the next e-meeting, however if Companies feel Rapporteur missed an important aspect to discuss, please add it below.

**Q12: Do you see another important aspect/topic to discuss in this email discussion?**

|  |  |
| --- | --- |
| **Company** | **Comments** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

3. Conclusion

TBD

4. Reference

1. RAN2#113-e meeting, meeting report;
2. R2-2102074, Offline on RAN enhancements QoS [AT113-e][506]; Nokia
3. 3GPP TS22.104, Service requirements for cyber-physical control applications in vertical domains, V17.4.0
4. TS23.501, System architecture for the 5G System (5GS), V17.0.0, 2021-03
5. R2-2102686, RAN Enhancement to support new QoS, Qualcomm Incorporated;
6. R2-2102726, Handling of Survival Time, CATT;
7. R2-2102993, RAN Enhancement for New QoS Parameters, Nokia, Nokia Shanghai Bell;
8. R2-2103060, RAN enhancements based on new QoS related parameters, Huawei, HiSilicon;
9. R2-2103125, Discussion on RAN enhancement to support survival time, vivo;
10. R2-2103196, Topics on new QoS handling, Fujitsu
11. R2-2103212, RAN enhancement based on new QoS, OPPO
12. R2-2103329, Further considerations on new QoS, ZTE Corporation, Sanechips, China Southern Power Grid Co., Ltd
13. R2-2103420, Discussion on RAN enhancements based on Survival Time, III
14. R2-2103429, RAN enhancements based on new QoS related parameters, Ericsson
15. R2-2103432, Entering and operating in the Survival Time state, Samsung Electronics GmbH
16. R2-2103689, Discussion on the RAN support for new QoS parameters, CMCC
17. R2-2103735, RAN2 Enhancements for Support of QoS Parameters, Intel Corporation
18. R2-2103798, Enhancements based on new QoS requirements, InterDigital
19. R2-2103896, Discussion on entering and exiting survival time state, Futurewei Technologies
20. R2-2104097, View on survival time mechanisms, LG Electronics UK
21. R2-2104225, Clarification on the survival time, Xiaomi Communications
22. R2-2104265, RAN enhancements based on new QoS, TCL Communication Ltd