e3GPP RAN WG2 Meeting #112e R2-2010764

November 2nd – 13th, 2020

Agenda Item: 8.10.2.1

Source: InterDigital (email discussion Rapporteur)

Title: [DRAFT] Summary of [AT112-e][103][NTN] RACH and HARQ feedback aspects

Document for: Discussion, Decision

# Introduction

This discussion document is intended to enable continuation of user plane discussions from RAN2#112e, specifically relating to RACH and a subset of HARQ feedback-related aspects:

** [AT112-e][103][NTN] RACH and HARQ feedback aspects (IDC)**

      Scope: Discuss (a revision of) p2, p3, p5, p10, p12, p9, p13 from [R2-2010455](file:///C:\Data\3GPP\Extracts\R2-2010455%20(R17%20NTN%20WI%20AI%208.10.2.1%20Summary%20of%20%5bPost111-e%5d%5b908%5d%5bNTN%5d).docx)

Intended outcome: summary of the offline discussion with e.g.:

* List of proposals for agreement (if any)
* List of proposals that require online discussions

Please note the following deadlines for company feedback have been provided by the Chair:

* Initial deadline (for companies' feedback): Monday 2020-11-09 17:00 UTC
* Initial deadline (for rapporteur's summary in R2-2010764):  Monday 2020-11-09 23:00 UTC

The following was also noted:

Proposals marked "for agreement" in R2-2010764 not challenged until Tuesday 2020-11-10 12:00 UTC will be declared as agreed by the session chair. For the rest the discussion will continue online.

# Discussion

## Offset calculation (P2/P3)

From previous discussion in [1] the following was proposed:

*Proposal 3: From RAN2 perspective, for UE with pre-compensation capability, start of the ra-ResponseWindow and msgB-ResponseWindow is offset by UE-gNB RTD in LEO/GEO.*

Companies which did not agree expressed concerns about the accuracy of the UE timing precompensation estimate, stating that the UE may start monitoring too early/late if the estimate is inaccurate or if the timing advance had changed (e.g. due to satellite movement in LEO). However, based on agreements from the previous meeting, RAN1 is discussing a TA margin to compensate for such sources of inaccuracy (however more details may be needed from RAN1) [2]:

* *FFS: The TA margin, if needed and indicated by the network (in order to account for the TA estimation uncertainty)*

An alternative solution was proposed in [3], where if the timing reference is at the gNB the offset of the start of ra-ResponseWindow (msgB-ResponseWindow) can be made in the first PDCCH occasion after the downlink symbol that has the same symbol number, slot number and system frame number as the last uplink symbol of the PRACH occasion where msg1(MsgA) was transmitted (companies can view a detailed description in the original proposing contribution).

**Question 1: For UE with pre-compensation capability, from RAN2 perspective which option(s) do you support to offset the start of the *ra-ResponseWindow* and *msgB-ResponseWindow*? (FFS RAN1 details on TA margin included in UE-gNB RTT estimate)**

**Option 1: UE-gNB RTT estimate used for Msg1/MsgA transmission;**

**Option 2: Based on DL timing (e.g. downlink symbol that has the same symbol number, slot number and system frame number as the last uplink symbol of the PRACH occasion where msg1/MsgA was transmitted);**

**Option 3: Other (please describe).**

|  |  |  |
| --- | --- | --- |
| **Company** | **Preferred Option(s)** | **Additional comments** |
| MediaTek | Option 1 | We prefer Option 1, i.e. UE-gNB RTT estimate |
| LG | Option 1 | Option 1 should be commonly used in NTN, i.e., for a UE with/without precompensation capability. |
| ZTE | Option 1 | And the details of RTT is pending on RAN1’s decision. No need to duplicated the discussion in RAN2 unless there are further RAN2 impacts identified. |
| Spreadtrum | Option 2 | There are RA-RNTI collisions caused by error of UE-gNB RTD estimation for option 1. |
| OPPO | Option 1 | We think UE can determine the offset for the start of the *ra-ResponseWindow* and *msgB-ResponseWindow* based on estimated UE-gNB RTT. TA margin might be one of potential approaches to ensure the accuracy. |
|  |  |  |
|  |  |  |
|  |  |  |

Another proposal addresses the ra-ContentionResolutionTimer:

*Proposal 2: For UE with pre-compensation capability, start of the ra-ContentionResolutionTimer is offset by UE-gNB RTD in LEO/GEO.*

A similar solution can be employed as above (i.e. UE-gNB estimate + TA margin), however the UE-calculated offset for the ra-ContentionResolutionTimer may additionally be refined in Msg2 by a TA. Alternatively, [3] proposes that the offset the start of the ra-ContentionResolutionTimer by starting it in the downlink symbol that has the same symbol number, slot number and system frame number as the first uplink symbol after the end of the Msg3 transmission.

**Question 2: For UE with pre-compensation capability, which option(s) do you support to offset the start of the *ra-ContentionResolutionTimer*? (FFS RAN1 details on TA margin included in UE-gNB RTT estimate)**

**Option 1: UE-gNB RTT estimate used for Msg1;**

**Option 2: UE-gNB RTT estimate used for Msg1 transmission corrected by TA in Msg2;**

**Option 3: Based on DL timing (e.g. by starting it in the downlink symbol that has the same symbol number, slot number and system frame number as the first uplink symbol after the end of the Msg3 transmission)**

**Option 4: Other (please describe).**

|  |  |  |
| --- | --- | --- |
| **Company** | **Preferred Option(s)** | **Additional comments** |
| MediaTek | Option 1/Option 2 |  |
| LG | Option 2 | Option 2 should be commonly used in NTN, i.e., for a UE with/without precompensation capability. |
| ZTE | Option 2 | After TA correction,  NW will based on this corrected TA to schedule Msg4 therefore it is preferred to use corrected TA as the offset. Also the corrected TA is more accurate which is beneficial for further power saving. |
| Spreadtrum | Option 2 |  |
| OPPO | Option 2 | We prefer Option 2. The UE-gNB RTT estimated by UE is refined in Msg2, hence the absolute UE-gNB RTT can be used to offset the start of the *ra-ContentionResolutionTimer*. |
|  |  |  |
|  |  |  |
|  |  |  |

## Extention of ra-ResponseWindow/msgB-ResponseWindow (P5)

The following proposal received very large majority support in the previous email discussion:

*Proposal 5: If the start of the ra-ResponseWindow and msgB-ResponseWindow is accurately compensated by UE-gNB RTD, ra-ResponseWindow and msgB-ResponseWindow are not extended in LEO/GEO.*

Note: it is assumed most companies still agree with this proposal (i.e. the 24/26). If a company does not reply to the below question it is assumed satisfactory.

**Question 3: For companies which *disagree* with the following proposal, please justify why an extension is necessary or an acceptable alternate wording:**

***If the start of the ra-ResponseWindow and msgB-ResponseWindow is accurately compensated by UE-gNB RTT, ra-ResponseWindow and msgB-ResponseWindow are not extended in LEO/GEO.***

|  |  |
| --- | --- |
| **Company** | **Justification for extension/alternate wording?** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Enabling/Disabling HARQ UL retransmission (P10/P12)

From previous discussion in [1], the following was proposed with comments focused primarily on the definition of what disabling HARQ uplink retransmission means:

*Proposal 10: From a RAN2 perspective, HARQ uplink retransmission relying on the decoding result of previous PUSCH transmission at the UE transmitter can be enabled/disabled in Rel-17 NTN (i.e. blind retransmission and slot aggregation if configured are not disabled).*

According to legacy procedure, a dynamic UL grant is associated with a HARQ process ID (PID), where upon initial PUSCH transmission the UE stores the TB in the HARQ buffer of the identified process.

If the UE receives a subsequent grant with the same HARQ PID and NDI toggled, the UE flushes the HARQ buffer and can perform LCP procedure to build a new TB.

If the TB was not successfully decoded, the gNB may provide a grant with the same HARQ PID and NDI *not* toggled, requiring the UE to retrieve the stored TB from the HARQ buffer of the identified process and perform a retansmission. The delay between initial TB transmission to reception of a retransmission grant is at least one RTT.

According to legacy behaviour, during this time new data is not assigned to this HARQ process as the NDI is not toggled. Due to large propagation delay in NTN, if enough TBs are unsuccessfully decoded (resulting in the associated HARQ PIDs not being reusable for a long time) HARQ stalling may occur.

In one interpretation of the current proposal, if HARQ uplink retransmission relying on the decoding result of previous PUSCH transmission at the UE transmitter is disabled for a HARQ process, the UE is not expected to receive a retransmission grant for that HARQ process. It is assumed this interpretation would result in the following behaviour:

* If not configured with any other retransmission mechanisms (e.g. slot aggregation, blind decoding etc.) there would be no need to store the TB in the HARQ buffer of the identified process;
* If the UE receives a grant for that HARQ PID with NDI *not* toggled, it would ignore the grant or assume a subsequent grant for that HARQ PID implicitely has NDI toggled;
* Retransmission would rely on higher layers (e.g. RLC) (or some other form of retransmission mechanism such as blind decoding).

An alternative interpretation is if the gNB did not successfully decode the TB, it may still send a grant with toggled NDI to allow new data for this HARQ PID. If the gNB intends to toggle the NDI regardless of the decoding result, there is no need to wait at least an RTT to assign a grant to the same HARQ PID.

According to RAN1 TS 38.214, as long as the transmission of the last PUSCH for the HARQ process is completed, the gNB may assign a grant to the same PID:

*The UE is not expected to be scheduled to transmit another PUSCH by DCI format 0\_0, 0\_1 or 0\_2 scrambled by C-RNTI or MCS-C-RNTI for a given HARQ process until after the end of the expected transmission of the last PUSCH for that HARQ process.*

As the duration of the PUSCH transmission is much less than an NTN RTT, the gNB can assign a HARQ PID to new data much faster and avoid HARQ stalling. Referring to P12, there was large majority support that the criteria/decision to enable/disable HARQ uplink retransmission is already under network control:

*Proposal 12: From RAN2 perspective, HARQ uplink retransmission at the UE transmitter can be enabled/disabled, but HARQ processes remain configured. The criteria to enable/disable HARQ uplink retransmission is under network control, and is signalled to UE via RRC in a semi-static manner.*

Therefore, if the network wants to disable HARQ uplink retransmission for a HARQ process it can simply provide a grant with NDI toggled without waiting for the decoding result of the previous PUSCH transmission. If the gNB wants to enable HARQ uplink retransmission, it can simply use legacy behaviour. It is noted that further discussion may be required to properly configure the value of *drx-HARQ-RTT-TimerUL*.

In summary, disabling HARQ uplink retransmission relying on the decoding result of previous PUSCH transmission at the UE transmitter can be interpreted in two ways:

1. *gNB can send grant with NDI toggled without waiting for decoding result or previous PUSCH transmission (i.e. up to gNB implementation):*

* Relies on legacy mechanisms and gNB implementation (i.e. no spec impact);
* HARQ uplink retransmission can be enabled/disabled dynamically based on NDI;
* Can send a grant with same HARQ process much faster than an NTN RTT;
* Does not place restrictions on scheduling.

1. *UE assumes it will not receive a retransmission grant based on gNB decoding result of previous PUSCH transmission:*

* If not configured with any other retransmission mechanisms (e.g. slot aggregation, blind decoding etc.) there would be no need to store the TB in the HARQ buffer of the identified process as retransmission is not expected (i.e. change of legacy behaviour);
* Retransmission relies on higher layers (e.g. RLC) or some other mechanism such as blind decoding or slot aggregation;
* If a UE receives a grant for a HARQ PID with NDI *not* toggled, it would ignore the grant or assume NDI is implicitely toggled (i.e. change of legacy behavior);
* Places restrictions on scheduling.

**Question 4: Which option do you support to enable/disable HARQ uplink retransmission relying on the decoding result of previous PUSCH transmission at the UE transmitter?:**

**Option 1: gNB can send grant with NDI toggled without waiting for decoding result of previous PUSCH transmission (i.e. up to gNB implementation);**

**Option 2: UE assumes it will not receive a retransmission grant based on gNB decoding result of previous PUSCH transmission;**

**Option 3: Other (please describe).**

|  |  |  |
| --- | --- | --- |
| **Company** | **Preferred Option** | **Additional comments** |
| MediaTek | Option 1 | We prefer Option 1 as it has no spec impact. |
| LG | Option 2 |  |
| ZTE | Option 1 (i.e., reusing legacy behavior) | We prefer to reusing the legacy behavior which has less specs impact and can guarantee both the transmission reliability and provide more flexibility while achieving the same purpose option 2 aims for. Option 2 propose a semi-static way to disabling the HARQ retransmission which might demage the resource efficiency since it is impossible to dynamically switch between the blind retransmission and normal HARQ retransmission. Also only relaying on RLC retransmission will lead to extra latency especially when radio quality becomes worse suddenly, e.g., due to blockage of obstacles. |
| Spreadtrum | Option 3 | For HARQ retransmission enabled processe based on PUSCH decoding result, TA is used as the offset to be applied to drx-HARQ-RTT-TimerUL value range. UE will only receive grant of the same HP ID with NDI toggled or not, after drx-HARQ-RTT-TimerUL expiration.  For HARQ retransmission disabled processe based on PUSCH decoding result, drx-HARQ-RTT-TimerUL is 0. UE expect to receive grant of the same HP ID with NDI toggled or not, from now on. |
| OPPO | Option 1 | We think disabling HARQ uplink retransmission is not to disable the entire UL HARQ retransmission, and retransmissions of TB in a bundle or based on blind scheduling could still be supported.  For Option 1, perhaps it should be “with NDI not being toggled”, if retransmission is the intention here. |
|  |  |  |
|  |  |  |
|  |  |  |

## Other aspects (P9/P13)

**Question 5: Which of the following aspects should be further studied in NTN?:**

1. **Report UE-calculated TA in e.g. msg3/msg5/msgA;**
2. **Enhancements to RSRP-based selection mechanism of 2-step vs. 4-step RACH;**
3. **Introduction of K\_offset in SI (pending RAN1 agreements).**
4. **LCP impact caused by disabling HARQ UL retransmission.**

|  |  |  |
| --- | --- | --- |
| **Company** | **Supported Option(s)** | **Additional comments** |
| MediaTek | Option 1, Option 3, Option 4. | Reporting of UE-calculated TA, K\_offset and LCP impacts on disabling HARQ UL retransmissions should be studied. |
| LG | Option 2 and 4 |  |
| ZTE | Option 1 | **For option 2:**  We understand the intention of RSRP-based RA type selection is to ensure the 2-step RACH can only be selected in case the radio condition is qualified, which provide flexibility on NW side to deploy the 2-step RACH with a good resource efficiency (i.e. only take the UE with good radio condition into account and leave the UE in the edge of cell to 4-step RACH). In NTN, we don’t see new requirement on RA type selection, thus we think the current mechanism can be reused and no optimization is needed on this aspect from RAN2’s perspective.  **For Option 3:**  Since RAN1 is still discussing the details on K\_offset (e.g., cell or beam specific, explicit or implicit indication, and etc.,), what’s need to be broadcast is still uncertain, it is preferred to postpone the discussion until RAN1 make more progress.  **For Option 4:**  We understand the intention of LCP enhancement is to distinguish the grant with fast retransmission (e.g. blind retransmission) and the grant with slow retransmission (e.g. HARQ retransmission) , and ensure the LCH with strict requirement on latency can be mapped to a grant with fast retransmission. However, we think similar requirement have been discussed before in IIOT and the allowedPHY-PriorityIndex has been introduced for the similar purpose. With the PHY-priority index, UL MAC SDUs from the logical channel can only be mapped to the dynamic grants indicating PHY-priority index equal to the values configured for that LCH. In NTN, the NW can simply configure the grant with/without HARQ with different PHY-priority index. No further optimization is needed. |
| Spreadtrum | 1, 2, 3, 4 |  |
| OPPO | Option 1/2/3/4 | We think all these aspects are worth being studied in the WI. |
|  |  |  |
|  |  |  |
|  |  |  |

# Summary

<to be completed pending company input>

# Conclusions

Based on company feedback, the following are proposed:

<to be completed pending company input>

# References

1. R2-2010455 – Summary of [Post111-e][908][NTN] RACH and HARQ feedback aspects – InterDigital
2. Chairman’s Notes RAN1#102-e 8.4 v004 – RAN1 Vice Chair
3. R2-2010980 – On Random Access in NTN - Ericsson
4. TS 38.321 – Medium Access Control (MAC) protocol specification v16.1.0