**3GPP TSG-RAN WG2 Meeting #112 electronic R2-200xxxx**

**Elbonia, Nov 2nd – 13th 2020**

**Agenda item:** 8.11.2

**Source:** Intel Corporation

**Title:** Report of [Post111-e][625][POS] End-to-end latency analysis (Intel)

**Document for:**  Discussion and decision

# Introduction

This contribution provides report for RAN WG2 email discussion:

* [Post111-e][625][POS] End-to-end latency analysis (Intel)

Scope: Discuss which nodes and which procedures are involved in a positioning latency analysis, and capture expected latency values where possible.

Intended outcome: Report to next meeting

Deadline: Oct 15th , 2020

Rapporteur proposes to divide the discussion in two phases:

**Phase 1**: Based on the contributions in last meeting, and the LS/agreements from RAN1, discuss which nodes, procedures, solution (UE based/UE assisted) and RRC state should be considered in our analysis, the basic procedure for positioning methods and latency assumption for different nodes, etc. The goal of this phase is to provide analysis for the Rel.16 RAT dependent NR positioning solutions.

Deadline: Sep 30th

**Phase 2**: Finalize the E2E latency value range for different positioning methods based on discussion in phase 1, and collect potential enhancements/directions to reduce the latency (companies should show the latency analysis when propose the enhancements).

Deadline: Oct 15th

To make it easier to find the correct contact delegate in each company for potential follow-up questions, the rapporteur encourages the delegates who provide input to provide their contact information in this table:

|  |  |
| --- | --- |
| Company | Delegate contact |
| COMPANY\_NAME | NAME ([email@address.com](mailto:email@address.com)) |
| Intel | yi.guo@intel.com |
| Ericsson | [ritesh.shreevastav@ericsson.com](mailto:ritesh.shreevastav@ericsson.com), fredrik.gunnarsson@ericsson.com |
| Qualcomm | Sven Fischer (sfischer@qti.qualcomm.com) |
|  |  |

# Discussion

## Phase 1 (Analysis for Rel.16 NR Positioning)

As indicated in [2],

|  |
| --- |
| *RAN WG1 evaluates physical layer latency and its potential reduction for NR Rel-17 positioning solutions. In order to evaluate End-To-End latency of NR positioning solutions the input from RAN WG2 is needed on latency components of NR/NG-RAN/5GC higher layer positioning protocols.*  *RAN WG1 respectfully asks if RAN WG2 can provide a list of latency components with corresponding range of values for the existing and any potential enhanced NR positioning solutions* |

Based on RAN1 LS, RAN2 should focus on the latency analysis of NR/NG-RAN/5GC higher layer positioning protocols, and RAN1 is responsible for physical layer latency analysis.

Below MO-LR/MT-LR procedures are cited from [8], TS23.273.



Figure 6.1.2-1: 5GC-MT-LR Procedure for the commercial location services



Figure 6.2-1: 5GC-MO-LR Procedure

Based on RAN1 agreements below, seems RAN1 will do the E2E latency analysis based on positioning procedure.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Agreement:   * Physical Layer Latency Start and End times are defined as follows:  |  |  |  | | --- | --- | --- | | **Method** | **Start** | **End** | | UE assisted DL-only & DL-ECID & Multi-RTT | Transmission of the PDSCH from the gNB carrying the LPP Request Location Information message | Successful decoding of the PUSCH carrying the LPP Provide Location Information message | | UL-only method & UL ECID & Multi-RTT | Reception by the gNB of the NRPPa measurement request message | The transmission by the gNB of the NRPPa measurement response message | | UE-based | Transmission of the PDSCH from the gNB carrying the LPP Request Location Information if applicable, otherwise,   * Alt. 1: transmission of the PUSCH carrying the MG Request from the UE. * Alt. 2: Transmission of the PDSCH from the gNB carrying the LPP message containing the assistance data * Alt. 3: Start of the Reception of DL PRS   Note: Suggest to downselect this at the next meeting.  Note: The high layers latency components may be subject to adjustment for different alternatives. | Successful decoding of the PUSCH at gNB carrying the LPP Provide Location Information message if applicable, otherwise Calculation of Location Estimate at the UE | |

In addition, based on SID, the E2E latency should be analysed based on rel-16 positioning solutions, i.e. positioning delay, step 5 in MO-LR and step 12 in MT-LR. Location service delay is out of RAN scope. But it would be good to check companies’ view on this.

* 1. Evaluate the achievable positioning accuracy and latency with the Rel-16 positioning solutions in (I)IoT scenarios and identify any performance gaps. [RAN1]

**Question 1-1: For RAN2 latency analysis of Rel.16 solutions, should we only consider the latency of positioning procedure, i.e. step 5 in MO-LR/step 12 in MT-LR or the latency of location service, i.e. whole procedure in MO-LR/MT-LR?**

**Alt 1: the latency of positioning, i.e. step 5 in MO-LR/step 12 in MT-LR**

**Alt 2: the latency of location service, i.e. whole procedure in MO-LR/MT-LR**

|  |  |  |
| --- | --- | --- |
| **Company** | **Alt 1 or Alt 2** | **Remark** |
| Intel | Alt 1 | The latency of positioning, i.e. step 5 in MO-LR and step 12 in MT-LR is under RAN control. The location service is out of RAN scope, and therefore should not be used for E2E latency analysis. |
| Ericsson | Alt 1 | Yes, Nodes involving RAN procedures and protocols (LPP, RRC, MAC) should only be studied |
| Qualcomm | Modification of Alt-2 | Deferred MT-LR for periodic/triggered events as specified in TS 23.273 section 6.3 and 6.7 should be added to the analysis. We think most applications (e.g., (I)IoT scenarios) require regular (e.g., periodic) location reports to an client.  An end-to-end latency analysis needs to consider the time between the client request for location (for MR-LR or MO-LR) or the triggering of location (for deferred MT-LR) and the availability of the location at the client.  However, only the basic steps in the procedures (e.g., MT-LR, MO-LR) need to be considered:   * No inclusion of roaming. * For MT-LR, no inclusion of privacy check based on current location (i.e., steps 16-23). * For MO-LR, no transfer to 3rd party (i.e., steps 8-11).   The rational for an end-to-end analysis is that user case requirements from SA1 are end-to-end and detailed definitions are available in Rel-16 to enable such analysis. However, for any enhancements that could affect the 5GCN, SA2 and CT WGs could be asked to comment. |
|  |  |  |

The positioning procedure involves multiple nodes, e.g. UE, NG-RAN, AMF, LMF, VGMLC, HGMLC, UDM, external Client, NEF, AF, etc.

[4] mentioned, “*procedures external to the PLMN (between LCS client and GMLC), procedures related to roaming case (between V-PLMN and H-PLMN), and privacy verification procedures (e.g. between GMLC and UDM) are not included in the latency evaluation*”.

[5] also mentioned “*Considering the latency of the procedures between LCS client, GMLC, AMF for triggering the positioning is tightly related to deployment and out of RAN scope. The delay analysis in this contribution does not consider these procedures, and only count the procedures between UE, LMF, AMF and gNB*”.

The question is whether all involved nodes should be considered in latency analysis.

**Question 1-2: For RAN2 latency analysis of Rel.16 solutions, which nodes shall be considered? UE, NG-RAN, AMF, LMF, VGMLC, HGMLC, UDM, external Client, NEF, AF.**

|  |  |  |
| --- | --- | --- |
| **Company** | **UE, NG-RAN, AMF, LMF, VGMLC, HGMLC, UDM, external Client, NEF, AF** | **Remark** |
| Intel | UE, NG-RAN, AMF and LMF | The latency for other nodes, e.g. GMLC, etc are out of RAN WG2 scope. |
| Ericsson | Protocols and Procedures involving these entities:  UE, NG-RAN, AMF, LMF | The latency for other nodes is out of RAN2 scope. |
| Qualcomm | UE, gNB, AMF, LMF, GMLC, UDM, client/NEF/AF | For the latency of nodes considered out of RAN2 scope above, we can assume the same numbers as for other CN interfaces (e.g., AMF-LMF signalling). As previously commented, we should align the analysis with SA1 requirements in order to determine whether they can be supported. |
|  |  |  |

**Question 1-3: If the answer of Question 1-1 is alt 2, for RAN2 latency analysis of Rel.16 solutions, should MO-LR, NI-LR, MT-LR all be considered? Or only some of them.**

|  |  |  |
| --- | --- | --- |
| **Company** | **MO-LR, NI-LR, MT-LR** | **Remark** |
| Intel | MO-LR and MT-LR | If companies would like to consider Alt 2, MO-LR and MT-LR should be considered since MO-LR and MT-LR are typical cases for commercial use cases. |
| Qualcomm | Deferred MT-LR for periodic/triggered events, basic MT-LR and MO-LR | We consider Deferred MT-LR more typical for low-latency use cases. In fact, latency for a single location MT-LR or MO-LR may be high or very high in practice due to overheads in interfacing with an external client and paging the UE for an MT-LR. |
|  |  |  |
|  |  |  |

As indicated in [1], RAN1 agreed:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Agreement:   * At least the following information is provided for positioning physical layer latency analysis:   + Source initiating request for positioning measurements/location for a given UE (UE, Network)   + Destination awaiting for positioning measurements/location for a given UE (UE, Network)   + Start and end triggers/events for physical layer latency evaluation     - For Rel.16 solutions, it is based on specification for each solution   + Initial and final RRC State of positioned UE (RRC IDLE, INACTIVE, CONNECTED) at the start and end time for the physical layer latency evaluation   + Positioning     - technique (enumeration): (1) DL-TDOA, (2) DL AoD, (3) UL-TDoA, (4) UL-AoA, (5) Multi-RTT, (6) E-CID     - type: DL, UL, DL+UL     - mode: UE-based, UE-assisted   + Latency component w/ value range and description, including information on any parallel (simultaneous) components   + Total latency value * Latency components are recommended to be captured in table and ordered consequently in time starting from the earliest one:  |  |  |  | | --- | --- | --- | | **Source [UE, NW]/Destination [UE, NW]**  **Positioning technique [DL-TDOA, E-CID, …], type [DL, UL, DL+UL], mode [UE-A, UE-B],**  **Initial and Final RRC States [IDLE, INACTIVE, CONNECTED]** | | | | **Latency Component** | **Value Range** | **Description of Latency Component** | | Start trigger |  |  | | Name of component 1 |  |  | | Name of component 2 |  |  | |  |  |  | | Name of last component |  |  | | End trigger |  |  | | Total values |  |  | |

In summary, for RAN1 physical latency analysis,

* All RRC state shall be considered (IDLE, INACTIVE, CONNECTED);
* All Rel-16 RAT dependent positioning methods shall be considered ((1) DL-TDOA, (2) DL AoD, (3) UL-TDoA, (4) UL-AoA, (5) Multi-RTT, (6) NR E-CID);
* All positioning types shall be considered (type: DL, UL, DL+UL);
* Both UE-based, UE-assisted shall be considered;
* List latency component w/ value range and description, including information on any parallel (simultaneous) components

Rapporteur’s comments, from RAN2 perspective:

* The positioning type (UL, DL, DL+UL) has been reflected by positioning techniques, therefore we do not need to list them separately;

**Question 1-4: Based on RAN1 agreements, shall both UE based and UE assisted be considered in RAN2 latency analysis? Or one of them ? If one of them, pls indicate which positioning methods.**

|  |  |  |
| --- | --- | --- |
| **Company** | **Both , UE based only, UE assisted only** | **Remark** |
| Intel | Both |  |
| Ericsson | Both |  |
| Qualcomm | Both | For the Rel-16 methods, the "positioning signalling" is the same for UE-based and UE-assisted (e.g., Step 10 in Figure 1 below may be performed by the UE or by the LMF). Differences occur based on the location of the consumer of the position estimate (i.e., UE or external client). |
|  |  |  |

**Question 1-5: For RRC state, should we consider both scenarios for RAN2 latency analysis? Or only one of them?**

* + Scenario 1: To support MO-LR, MT-LR, NI-LR procedure, if the UE was in IDLE/INACTIVE, the state transmission to CONNECTED is needed before the positioning procedure (step 5 for MO-LR and step 12 for MT-LR) in order to exchange the message between the UE and the network.
  + Note: It is only applicable for alt 2 in question 1-1. For alt1, the transition time is not counted in positioning procedure itself although it is part of location service procedure.
  + Scenario 2: For the scenario that the UE obtains the assistance data via system information and provides the results to application inside the UE. The UE does not need to exchange the message with network, and the state transmission is not needed, i.e. initiate state could be IDLE, INACTIVE or CONNECTED, and the UE does not change the state when performing the positioning.
  + Note 1: this is based on rel-16, i.e. the UE cannot send/receive the data, reference signals in IDLE or INACTIVE mode.

|  |  |  |
| --- | --- | --- |
| **Company** | **Both, scenario 1 only or scenario 2 only** | **Remark** |
| Intel | Scenario 2 only | Scenario 1 and 2 are already supported in Rel-16. But procedure of scenario 1 is not part of step 5 for MO-LR and step 12 for MT-LR.  The scenario 1 could be considered only if the whole procedure for location service is considered, i.e. alt2 in question 1-1. |
| Ericsson | Scenario 2 including both UE-A/UE-B cases | For positioning latency there are 2 definitions:  a) TTFF: For every positioning method there is TTFF. Scenario 1 includes that and SA2, RAN3 are also part of this.  b) After TTFF is achieved; what is the requirement in terms of interval/frequency of providing the positioning to end user/client. (for example, once every 20ms).  Hence, from RAN2 perspective mainly the b) should be taken into consideration; i.e once the UE is in the ready state for performing positioning procedure; then the latency involved as below   * Obtaining necessary AD * Performing measurement * Reporting the measurement (UE assisted) * Positioning computation * Providing the position to UE (if client is in UE)   For GNSS based positioning method, TTFF depends upon whether UE has recent almanac/ephemeris data available or not. If it is then GNSS positioning procedure runs fast.  Similarly, for RAT dependent to run DL-TDOA; NR-ECID procedure is a prerequisite procedure hence this procedure is part of TTFF.  The latency aspect of TTFF that RAN2 deals with for example NR-ECID/Capability exchange can also be evaluated. However, the real aim should be to check what is the delay involved in b) i.e, After TTFF |
| Qualcomm |  | The scenarios are unclear; i.e., it is not clear why Scenario 2 considers assistance data delivery only. In any case, the UE needs to have a NAS signalling connection (i.e., CM connected state) and will be in RRC connected state after Step 6 and step 1 for MT-LR and MO-LR, respectively. |
|  |  |  |

Refer to TS38.305, the procedures for different positioning methods are listed as below (**the details of step 5 for MO-LR and step 12 for MT-LR**).

**1 DL-TDOA/DL-AoD**



**Figure 1 procedure for DL-TDOA/DL-AoD**

**Question 1-6: Any comments on DL-TDOA/DL-AoD procedure? Can we use it for E2E latency analysis for DL-TDOA/DL-AoD?**

**Note 1: The procedure involves other nodes, e.g. GMLC is not shown since the procedure is focusing on the details of step 5 for MO-LR and step 12 for MT-LR;**

**Note 2: The transition from IDLE/INACTIVE to CONNECTED mode is not shown in the figure;**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Remark** |
| Intel | Yes |  |
| Ericsson | Another alternative also possible | When Client is in UE  For UE based, it is possible that UE can compute positioning using these methods without having to have LPP transaction. This will minimize the latency.  UE can obtain ciphering keys. NW may broadcast the relevant AD. UE can perform measurement and consume the positioning location without having to send any LPP message. This is already supported by standard, but we can discuss if any companies think it is unsupported. |
| Qualcomm | With modification | For baseline, Step 4 should not be needed. We can assume that the LMF obeys the UE capabilities and provides the required assistance data (i.e., assistanceAvailability=FALSE). In addition, step 5 would then typically occur before step 3. For a deferred MT-LR, steps 1 and 2 are not needed as an LMF only needs to obtain the UE capabilities once. Additional optimization is also possible for a deferred MT-LR (e.g. as defined in TS 23.273 clause 6.7.1), where step 3 does not occur. This is one reason to include deferred MT-LR in the evaluation – since it has the potential for lower latency than a normal MT-LR, MO-LR or NI-LR. |
|  |  |  |

**2 UL-TDOA/UL-AoA**



**Figure 2 procedure for UL-TDOA/UL-AoA**

**Question 1-7: Any comments on UL-TDOA/UL-AoA procedure? Can we use it for E2E latency analysis for UL-TDOA/UL-AoA?**

**Note 1: The procedure involves other nodes, e.g. GMLC is not shown since the procedure is focusing on the details of step 5 for MO-LR and step 12 for MT-LR;**

**Note 2: The transition from IDLE/INACTIVE to CONNECTED mode is not shown in the figure;**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Remark** |
| Intel | Yes |  |
| Ericsson | Yes |  |
| Qualcomm | With modification | NRPPa Positioning Activation has also a response message.  NRPPa Positioning Deactivation could be added to the end of the procedure. |
|  |  |  |

**3 Multi-RTT**



**Figure 3 procedure for Multi-RTT**

**Question 1-8: Any comments on Multi-RTT procedure? Can we use it for E2E latency analysis for Multi-RTT?**

**Note 1: The procedure involves other nodes, e.g. GMLC is not shown since the procedure is focusing on the details of step 5 for MO-LR and step 12 for MT-LR;**

**Note 2: The transition from IDLE/INACTIVE to CONNECTED mode is not shown in the figure;**

**Note 3: in the figure, we did not distinguish serving gNB and measured gNB and did not show multiple measured gNBs although they could be different.**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Remark** |
| Intel | Yes |  |
| Ericsson | yes |  |
| Qualcomm | With modification | NRPPa Positioning Activation has also a response message.  NRPPa Positioning Deactivation could be added to the end of the procedure.  Step 9 and 15 need to happen concurrently.  For baseline, Step 11 (Request Assistance Data should not be needed (see our response to Question 1-6). For a deferred MT-LR, steps 1 and 2 can also be avoided. |
|  |  |  |

**4 NR E-CID**



**Figure 4-1 procedure for Downlink NR E-CID**

**Question 1-9: Any comments on Downlink NR E-CID procedure? Can we use it for E2E latency analysis for Downlink NR E-CID?**

**Note 1: The procedure involves other nodes, e.g. GMLC is not shown since the procedure is focusing on the details of step 5 for MO-LR and step 12 for MT-LR;**

**Note 2: The transition from IDLE/INACTIVE to CONNECTED mode is not shown in the figure;**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Remark** |
| Intel | Yes |  |
| Ericsson | Yes | Step 4 could be seen as optional, since DL E-CID is based on measurements that the UE has available. |
| Qualcomm |  | The assumptions for Step 4 require some clarification. According to TS 38.305:  "Although NR E-CID positioning may utilise some of the same measurements as the measurement control system in the RRC protocol, the UE generally is not expected to make additional measurements for the sole purpose of positioning; i.e., the positioning procedures do not supply a measurement configuration or measurement control message, and the UE reports the measurements that it has available rather than being required to take additional measurement actions." |
|  |  |  |



**Figure 4-2 procedure for Uplink NR E-CID**

**Question 1-10: Any comments on Uplink NR E-CID procedure? Can we use it for E2E latency analysis for Uplink NR E-CID?**

**Note 1: The procedure involves other nodes, e.g. GMLC is not shown since the procedure is focusing on the details of step 5 for MO-LR and step 12 for MT-LR;**

**Note 2: The transition from IDLE/INACTIVE to CONNECTED mode is not shown in the figure;**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Remark** |
| Intel | Yes |  |
| Ericsson | Yes | Step 3 could be seen as optional, since DL E-CID is based on measurements that the UE has available. |
| Qualcomm |  | Step 3 and/or Step 2 are unclear; UE seems not transmitting to enable gNB measurements. We assume the RRC Measurement Configuration (Step 2) is for RRM measurements, therefore, some SRS configuration/activation seems needed as well. Also, some UE measurement report seems missing. |
|  |  |  |

**Question 1-11: Based on RAN1 agreements, shall all RAT dependent positioning methods be considered in RAN2 latency analysis, ((1) DL-TDOA, (2) DL AoD, (3) UL-TDoA, (4) UL-AoA, (5) Multi-RTT, (6) NR E-CID including UL NR E-CID)? or only some of them? If some, pls indicate which positioning methods.**

|  |  |  |
| --- | --- | --- |
| **Company** | **All or some** | **Remark** |
| Intel | All |  |
| Ericsson | 1, 3, 5 | Angle based and ECID could be to some extend covered by 1,3, 5 |
| Qualcomm |  | We see no need for E-CID, since low-accuracy positioning should not require low-latency, but O.K. to include. |
|  |  |  |

Based on the procedures shown above, the procedure for DL-TDOA and DL-AoD is same, and forUL-AoA and UL-TDOA is same.

**Question 1-12: In RAN2 latency analysis, can DL-TDOA and DL AoD be handled together? And UL-TDoA and UL-AoA be handled together, i.e. in the same table?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Remark** |
| Intel | Y | From RAN2 perspective, there is no different although from physical layer perspective, the measurement delay could be different. |
| Ericsson | Y |  |
| Qualcomm | Yes |  |
|  |  |  |

Regarding the table used for collecting the latency analysis, RAN1 agreed below template:

|  |  |  |
| --- | --- | --- |
| **Source [UE, NW]/Destination [UE, NW]**  **Positioning technique [DL-TDOA, E-CID, …], type [DL, UL, DL+UL], mode [UE-A, UE-B],**  **Initial and Final RRC States [IDLE, INACTIVE, CONNECTED]** | | |
| **Latency Component** | **Value Range** | **Description of Latency Component** |
| Start trigger |  |  |
| Name of component 1 |  |  |
| Name of component 2 |  |  |
|  |  |  |
| Name of last component |  |  |
| End trigger |  |  |
| Total values |  |  |

Rapporteur’s comments:

* Source [UE, NW]/Destination [UE, NW] are not needed, it can be reflected in procedure steps if we show the procedure figure together with the table;
* Initial and Final RRC States [IDLE, INACTIVE, CONNECTED] is not needed for each procedure since it is common for all positioning methods, and could be captured in separate table, or just as separate steps in the table;

If we follow RAN1 table, and consider MO-LR, MT-LR, 5 positioning methods (same table for DL-TDOA/DL-AoD and same table for UL-TDOA/UL-AoD), 3 UE states, 2 modes, we need 60 tables.

If we have positioning methods specific table and common table, the total number will be 26 (20+6); In addition, if we only consider positioning procedure instead of location service procedure, then only table 1 below is needed, the total number will be 20;

Therefore Rapporteur would suggest to use below table to capture latency analysis in phase 2:

**Table 1: positioning method specific table**

|  |  |  |
| --- | --- | --- |
| **Positioning technique [DL-TDOA/DL-AoD, Downlink NR E-CID, …], mode [UE-A, UE-B (IDLE, INACTIVE, CONNECTED)],**  **Focusing on the latency caused by steps in figure 1-figure 4** | | |
| **Latency Component** | **Value Range** | **Description of Latency Component** |
| Step 1 |  |  |
| Step 2 |  |  |
| Step 3 |  |  |
| .. |  |  |
|  |  |  |
|  |  |  |
| Total values |  |  |

**Table 2: common procedure** **(rely on the answer of question 1-1, i.e. whether Alt2 is used for E2E latency analysis)**

|  |  |  |
| --- | --- | --- |
| **e.g. location request for MO-LR, MT-LR when the UE is in IDLE, INACTIVE or CONNECTED, and location response, etc.**  **Refer to the figure 6.2-1 for MO-LR (except step 5) and 6.1.2-1 for MT-LR (except step 12);** | | |
| **Latency Component** | **Value Range** | **Description of Latency Component** |
| Step 1 |  |  |
| Step 2 |  |  |
| Step 3 |  |  |
| .. |  |  |
|  |  |  |
|  |  |  |
| Total values |  |  |

**Question 1-13: Do companies agree to use table1 (positioning methods specific) and table 2 (common procedure) described above to capture the latency analysis results? Or any other suggestions?**

**Note: The need of table 2 relies on the answer of question 1-1, i.e. whether Alt2 is used for E2E latency analysis)**

|  |  |  |
| --- | --- | --- |
| **Company** | **Table 1 and Table 2, Table 1 only** | **Remark** |
| Intel | Table 1 | See above, we only need to consider step 5 for MO-LR and step 12 for MT-LR, and therefore table 1 is enough for RAN2 analysis. |
| Ericsson | Table 1 |  |
| Qualcomm |  | A simpler and more flexible approach is as follows.   * Construct a table T0 containing entries for atomic operations (e.g. LPP processing time, transmission + propagation time over a given link, location computation time, measurement time) and include labels and agree latency values. Labels might, for example, include those used for Q 1-14 below or could be more generic (e.g. O1, O2, O3 etc.). * Construct a single table T1 with entries for composite operations (e.g. LPP or NRPPa end to end message transfer time) which can belong to a positioning or location related procedure. Map each composite operation to constituent atomic operations and/or to other (simpler) composite operations and determine latency values and include labels. * Define and agree a message flow for each type of positioning procedure (DL, UL, UL-DL) and map each procedure to constituent atomic and composite operations used in each procedure. Determine the overall latency using tables T0 and T1. It is possible that variants of the positioning procedures will be needed for deferred MT-LR and certain enhancements of location solutions. Enter the overall latency for each positioning procedure in a table T2. * Define and agree a message flow for each location solution (MT-LR, MO-LR, NI-LR, deferred MT-LR) and map each message flow to a positioning procedure and atomic and composite operations. Determine the overall latency using tables T0, T1 and T2. Enter the overall latencies in a table T3 which can provide end to end latency. * Attempt to account for differences in UE initial/final state and UE based versus UE assisted mode and other such differences via differences in atomic and/or composite operations referenced from tables T0 and T1. This should lead to fixed differences in Table T3 – which can then assume some default configuration and include deltas for the differences in a Note.   The value of this approach is to better identify common operations and common positioning procedures used in different location solutions and ensure that the same latency assumptions are used. This may also reduce documentation (e.g. number and size of tables). As an example, a location solution S1 might list the constituent operations being used as O1, O2, O3, O4 etc. with the end to end latency being the sum of entries in tables T0 and T1 for these operations. An enhanced version of S1 can then list (e.g.) a subset of these operations (e.g. O1, O3, O4 etc. with O2 omitted). It would then be clear that O2 was omitted and that O1, O3 and O4 were common. |
|  |  |  |

In addition, it would be good to have common understanding on the latency assumption for:

* State transition: IDLE to CONNECTED, INACTIVE to CONNECTED;
* Processing delay:
  + UE RRC processing delay, UE LPP processing delay (capability transfer, assistance data transfer, location request), UE MAC processing delay;
  + gNB RRC processing delay, gNB NRPPa processing delay;
  + AMF processing delay;
  + LMF processing delay;
* transmission delay:
  + UE from/to gNB;
  + gNB from/to AMF;
  + AMF from/to LMF;
  + AMF from/to GMLC;

**Question 1-14: Companies are invited to provide the input on latency assumptions (Unit= ms)**

**Table 3: Idle/INACTIVE to CONNECTED, and UE processing time on RRC, LPP and MAC (Unit= ms);**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Company** | **TIdle-conn** | **TInac-conn** | **TUE-RRCReconf** | **TUE-RRCDLInfo** | **TUE-RRCULInfo** | **TUE-RRCLocationMeas** | **TUE-LPPCapab** | **TUE-LPPAssi** | **TUE-LPPLocationRe** | **TUE-MAC-SRSAct** |
| **Intel** | **36.3-62.5** | **11.3-18.5 [9]** | **10 [10]** | **5** | **2** | **2** | **20-80 [10]** | **10** | **5** | **1** |
| **Ericsson** | **48** |  |  |  |  |  |  |  |  |  |

**Table 4: gNB processing time on RRC and NRPPa, AMF/LMF processing time, transmission delay between nodes, gNB measurement delay and LMF calculation delay (Unit= ms);**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Company** | **TgNB-RRC** | **TgNB-NRPPa** | **TgNB-NAS/LPP** | **TAMF** | **TLMF** | **TUE-gNB** | **TgNB-AMF** | **TAMF-LMF** | **TAMF-GMLC** | **TgNB-Meas** | **TLMF-Calc** |
| **Intel** | **3 [9]** | **3** | **3** | **3** | **3** | **0** | **3-10** | **3-10** | **3-10** | **RAN1** | **5** |
| **Ericsson** |  |  |  |  |  |  |  | **1-2** |  |  | **2** |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |
| --- | --- |
| **Company** | **Remark** |
| Intel | 1 **TIdle-conn** (INACTIVE to CONNECTED), referred to TS37.910 [9];  2 **TIdle-conn** (IDLE to CONNECTED), on top of **TInac-conn,** added   |  | | --- | | *RRCReconfig+SMC = T UE processing (10-15ms)*  *2 Initiate NAS = 2\*T gNB processing (2\*3ms) + TAMF processing + 2\*T-AMF-gNB* |   3 **TUE-LPPCapab** (LPP capability processing time) referred to RRC capability processing time in TS38.331 [10];  4 **TUE-LPPAssi** (LPP assistance data) referred to RRCReconfiguration processing time in TS38.331 [10];  5 **TgNB-RRC** referred to TS37.910 [9], *the processing delay in gNB (L2 and RRC) has been reduced to 3 ms*  6 **TgNB-NRPPa** and **TgNB-NAS/LPP** referred to TgNB-RRC  7 **TgNB-Meas** should be similar to **TUE-Meas,** and wait for RAN1**.**  8 The transmission delay is tightly related to backhaul situation; |
| Ericsson | An optimized AMF/LMF for latency can have 1-2ms delay/latency between AMF to LMF. For Non-Public Network where 5GS can be deployed locally (in factory premises) this delay/latency could further be lower. Location can be computed in location server every 2ms. |
| Qualcomm | |  |  |  |  | | --- | --- | --- | --- | | Delay Component | | Assumption  [ms] | Comment | | MAC-CE Processing | TMAC-CE | 3 | For SRS activation.  TS 38.214; the command goes into effect 3 ms after the UE sends Ack for the PDSCH carrying the transport block that included the MAC-CE. | | RRC Message Procesing | TUL-RRC  TDL-RRC | 5 | NAS UL/DL Transfer and Location Measurement Indication. Since these procedures are not listed in TS 38.331, section 12, the minimum value for the RRC procedures in Table 12.1-1 of 38.331 is assumed. | | Transition from RRC Inactive to Connected state | TRRCconnect | 15 | 11.3 ms – 18.5 ms dependent on configuration/deployment (TR 37.910). Assume an "average" value. | | RRC Reconfiguration | TRRC-ReConf | 10 | TS 38.331, section 12. | | gNB processing for determining measurement gaps | TgNB-Proc | 10 | Determining/scheduling appropriate measurement gaps as requested by the UE. | | LPP Message Processing | TLPP | 10 | Decoding/encoding and processing of the encapsulated LPP PDU (e.g., providing/ obtaining relevant parameter to/from physical layer).  The same is assumed for all LPP Procedures. | | NRPPa Message Processing | TNRPPa | 10 | Decoding/encoding and processing of the encapsulated NRPPa PDU (e.g., providing/ obtaining relevant parameter to/from physical layer).  The same is assumed for all NRPPa Procedures. | | DL-PRS/UL-SRS Measurement Time | TDL-Meas |  | Depends on RAN1 | | TUL-Meas | | Position Calculation | TPosCalc | 30 | Value could also be determined by RAN1 | | Core Network Signalling | TUL-NG-C  TDL-NG-C  TUL-NL1  TDL-NL1  TNlmf  TLCS-Resp | 3 | Transmission, propagation and processing time. Depends on network configuration; e.g. signaling link bandwidth values and distances between nodes.  On average, we assume the same value on all CN interfaces. | |
|  |  |

## Phase 2 discussion

### Latency analysis on Rel-16 RAT dependent positioning methods

### Potential solutions/directions on latency reduction

# Summary

To be added:

# Reference

[1] Chairman's Notes RAN1#102-e v022

[2] R1-2007264 LS on Latency of NR Positioning Protocols, RAN1

[3] R2-2006672 Discussion on ehancements for commercial use cases, CATT

[4] R2-2006578 Discussion on R17 positioning enhancement, Huawei, HiSilicon

[5] R2-2006750 Consideration on the support of low latency requirement, Intel Corporation

[6] R2-2007587 End-to-end latency reduction for DL/UL positioning, InterDigital, Inc.

[7] R2-2008261 [AT111-e][612][POS] Assumptions for analysis of commercial use cases (Ericsson) Ericsson

[8] TS 23.273, 5G System (5GS) Location Services (LCS); Stage 2.

[9] TS37.910

[10]TS38.331