**3GPP TSG RAN WG2 Meeting #121 R2-230xxxx  
Athens, Greece, 27th Feb.– 3rd March 2023**

**Agenda item: 8.16.2**

**Source: Apple**

**Title: Summary of [AT121][025]: Progress table of analyzing data collection framework (Apple)**

**WID/SID: FS\_NR\_AIML\_air – Release 18**

**Document for: Discussion and Decision**

# 1 Introduction

This is the summary of below offline discussion:

[R2-2300708](file:///C:\Users\johan\OneDrive\Dokument\3GPP\tsg_ran\WG2_RL2\RAN2\Docs\R2-2300708.zip) Open issues on AI/ML model delivery and data collection in post-meeting email discussion Apple discussion Rel-18 FS\_NR\_AIML\_air

- QC think we need the requirement.

* The table in this doc is endorsed as starting point
* Offline 025 (Apple) progress the table of methods and characteristics. Aim to endorse.

# 2 Discussion

The table 2 of R2-2300708 was endorsed as starting point of discussion. So, the table is copied below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Terminated entity | Allowed payload size | Report latency | Supported report type | Security and Privacy |
| Logged MDT | Between UE and TCE/OAM | <=64kbyte  (MDT buffer size limit) | Long  (Reported after entering CONNECTED) | Upon gNB request after entering CONNECTED | Security via RRC message,  Privacy via user consent |
| Immediate MDT | Between UE and TCE/OAM | <=9kbyte or 144kbyte  (with 16 segments) | Medium  (~20ms RRC signaling latency) | Event triggered report,  Periodic reporting | Security via RRC message,  Privacy via user consent |
| L3 measurements | Between UE and gNB | <=9kbyte or 144kbyte  (with 16 segments) | Medium  (~20ms RRC signaling latency) | Event triggered report,  Periodic reporting | Security via RRC message |
| L1 measurement (CSI reporting) | Between UE and gNB | Small  (<1706bit in PUCCH,  <3840bit in PUSCH) | Short  (can be symbol or slot level) | Aperiodic report,  Semi-persistent report,  Periodic report | No security |
| UAI | Between UE and gNB | <=9kbyte or 144kbyte  (with 16 segments) | Medium  (~20ms RRC signaling latency) | Up to UE implementation when to report | Security via RRC message |
| Early measurements | Between UE and gNB | <=9kbyte or 144kbyte  (with 16 segments) | Long  (Reported after entering CONNECTED) | Upon gNB request after entering CONNECTED | Security via RRC message |
| LPP | Between UE and LMF | <=64K payload  (NAS payload container limit) | Longer than L3 measurement  (Extra forward latency between LMF and gNB) | UE-triggered or NW-triggered | Security via RRC message |
| EVEX | Between UE and DCAF/ASP | Supports all file sizes | Longer than L3 measurements | UE triggered or network triggered | Security via user consent |

To make progress, Rapporteur would like to collect companies' view on below 2 questions. The table will be updated based on companies' input.

First, Rapporteur would like to collect companies' comments on reviewing the existing contents of the table.

**Q1: Companies are invited to share their comments on the existing contents of the table (Please do not insert comments or make trackable edits in above table, which will be hard for Rapporteur to track and respond your comments)**

|  |  |  |
| --- | --- | --- |
| **Company** | **Comments on existing contents of the table** | **Rapporteur response** |
| Intel | We think report latency may misleading it is the delay of the whole data collection framework. We prefer to change it into “signaling delay” to reflect it is the signaling processing delay, rather than collection, etc. More latency could also be considered, measurement duration, report interval, etc, as we raised in next response. |  |
| Interdigital | We agree with Intel’s comment above. |  |
| OPPO | 1. Allowed payload size column is misleading as the differentiation principle is not consistent, for any data collection framework using RRC signaling, the data limitation is the same, i.e. 9kbyte per RRC message, this is also applied to logged MDT and LPP data collection procedure, in theory, UE can report endless data via multiple RRC message if Periodic reporting or network sustaining request is applied;   As for L1 measurements, Allowed payload size is 1706bit per PUCCH message or 3840bit per PUSCH message;  As for the UE data buffer capability and message segmentation aspects, they are totally different topics, which can be decoupled at this stage, we can add more columns for them if needed;  So we suggest to change the column name from ‘Allowed payload size’ to ‘Allowed payload size per message’  The same issue is found for Report latency column, the differentiation principle is not consistent, for any data collection framework using RRC signaling, the Report latency is almost the same, i.e. ~20ms for RRC message, this is also applied to logged MDT and LPP data collection procedure, maybe consider end to end delay/(near)real-time or non-real time aspect makes more sense, we can clarify this. |  |
| Huawei, HiSilicon | We have some comments:  For the wording like “(with 16 segments)”, we suggest to clarify it to: “(with 16 RRC segments)”.  For report latency, we think the definition can be clarified, e.g. it refers to the latency between when the UE has got the data and when the UE transmits the data to the NW.  For security, “Security via RRC message” can be clarified, e.g. whether it refers to AS security activation, or also refers to PDCP ciphering/IP, or both. |  |
| Mediatek | We are supportive to consider the aspects listed in the table. But some aspects need to be clarified. For allowed payload size, currently UL DCCH message segmentation is only applicable to *UECapabilityInformation* and *MeasurementReportAppLayer*. Therefore, the allowed payload size should be 9kbyte for L3 measurement, UAI and early measurement. We can’t assume that UL DCCH message segmentation has been already supported for those UL messages.  For report latency, we may need to clarify from which time point the report latency starts. In this table, the report latency only considers the RRC message transmission and processing latency. However, what matters may be the latency rom availability of the measurement report at the UE side to the availability of the measurement report at the network side. I think this is also the reason why supported report type is considered. Taking L3 measurement for example, UE needs to wait for a while e.g., TTT and then sends the measurement report. If UL message segmentation is considered, the latency will be even longer.  For L1 measurement, the intention is L1 measurement is not security protected. For L3 measurement, LPP, UAI, etc., it is assumed that there is no privacy issue. If it is the case, we can capture this in the table. |  |
| Qualcomm | We want to add Event Exposer (EVEX) to the table. On the segmentation and report latency, we have a similar view. We also want to highlight report latency has different implications for data collection for different LCM purposes. Therefore, the reported latency should not be a determining factor for data collection for all LCM purposes. |  |
| Xiaomi | We understand the ‘report latency’ refers to the time between data collected until collected data is reported successfully, which includes both procedural delay and signaling delay. Maybe better to clarify that. |  |
| vivo | For the framework,   * SON is not included. * As we agreed that model delivery that serves the use cases in the SI is within RAN2 scope, the NRPPa between gNB and LMF can also be considered.   For the performance metric:   * For the terminated entity, the current content is misleading that gNB/LMF can provide data to gNB. We supporse it shall be ‘TCE, LMF or gNB’, rather than ‘between UE and xxx’ * For the allowed payload, LPP is transferred as NAS-PDU container in RRC message. So we preferred to split it into ‘allowed payload size per message’ and ‘allowed total payload with segments’ * For the latency, it can also split into ‘latency to generate the data’ and ‘latency to report the data after generated’   For the supported report type, it can also split into ‘record type’ and report type. |  |
| Spreadtrum | As for Allowed payload size, it should be the maximum payload per transmission rather than the total data size, e.g., for logged MDT, the successful reported data will be deleted from UE buffer, and then new collected data can be stored. Thus the total data size that can be reported may more than 64kbytes. And we tend to agree with OPPO, the maximum payload per transmission should be 9kbyte for those frameworks who use RRC message. |  |
| CMCC | For the definition of report latency, we share similar view with Xiaomi that it refers to the time between the UE has collected data and collected data is reported successfully. So we suggest to change ‘report latency’ to ‘data collection and reporting latency’, which would be more clear. |  |

Then, during online discussion, some companies suggested to add more columns for the table (i.e. other performance metrics of data collection framework). To make progress,Rapporteur would like to collect companies' view. Meanwhile please note that EVEX framework was agreed that it can be discussed next meeting because companies need more time to understand this framework.

* R2 may consider including the existing EVEX framework for this SI, FFS exactly what this means, can discuss next meeting.

Thus, **we will not discuss EVEX framework in this offline discussion**.

**Q2: Companies are invited to share their view whether / what other performance metrics can be added as new column(s) of the table.**

|  |  |  |
| --- | --- | --- |
| **Company** | **Please provide other performance metric(s) you think necessary. If any, please also provide your analysis of the new metric(s) for these data collection framework.** | **Rapporteur response** |
| Intel | we think it would be good to also consider other aspects to compare: 1) suitable data type, 2) measurement duration, 3) report interval, as those may reflect the potential latency of different data collection framework. Hence, we propose to merge below information in R2-2300418 with above table:   |  |  |  |  | | --- | --- | --- | --- | |  | Suitable data type | Measurement duration | Report interval | | Logged MDT | Non-urgent data, including location info | Min: 10min  Max: 120min | Min: 320ms  Max: infinity | | Immediate MDT | Min: 120ms  Max: 30min | Min: 120ms  Max: 30min | | RRM measurement reports | Real-time radio performance | | UE assistance information | Assistance information to show UE preference | Sec-level  Varies from 0s to 600s depends on the reported data | Configuration or upon change of problem information | | LPP Provide location information | Location information | immediate | Per request | | CSI reporting framework | CSI reporting | Slot-level  Min: 4 slot  Max: 320 slot | Slot-level  Min: 4 slot  Max: 320 slot | |  |
| Interdigital | Apart from the aspects indicated above by Intel, it would be good to elaborate more on the content of the data that is currently collected/reported in these frameworks. For example,  RRM measurements contain detailed cell level, beam level information that can be based on SSB or CSI-RS measurements, while logged measurements contain only cell level measurements.  Also, we can also include information about RRC states when the measurements/data are taken/logged. This is important because performing/logging the measurements in IDLE/INACTIVE will have limitation as the reason for putting in IDLE/INACTIVE is for power saving purposes. On the other hand, performing/logging measurements in CONNECTED may have a downside in terms of other requirements such as buffering size when compared to IDLE/INACTIVE. |  |
| OPPO | * As mentioned in Q1, whether UE data buffer is allowed or not can be considered as another domain, as you know, dataset can be large, data buffer may or may not be needed for UE dataset reporting; * UE applied RRC state during data collection can be considered also as some data collection framework only collect data from RRC idle and inactive, while others from RRC connected; * Data type, some framework collects measurement, others collects UE assistant info; * For message segmentation aspect, we can drop it now, as mentioned in Q1, in theory, UE can report endless data via multiple RRC message if Periodic reporting or network sustaining request is applied, we can consider this later; * Add a note to this table to say all the above data collection framework is using control plane signaling; |  |
| Huawei, HiSilicon | We have some comments:  We agree with the email rapporteur that EVEX framework is not discussed in this email discussion.  We observe that for logged MDT and early measurements, the UE logs information when it is in RRC idle/inactive states, while the UE logs information when it is in RRC connected state for other mechanisms. We think the RRC state where the UE does logging is important, so it is suggested to clarify it in the table, e.g. add a new column like “RRC state where the data is got/RRC state corresponding to collected data”.  Since the data collection framework will be anyway used in specific LCM in specific use cases, it may be good to also analyze them as well. However, we do not have strong views to add them for now, as we do think that we have listed quite a lot of frameworks here. So the LCM/use cases metrics could be discussed in future meetings. |  |
| Mediatek | We are open to consider more metrics in the future. Just as commented in Q1, the definition of report latency needs to be clarified. Intel provides another alternative on how to consider the overall latency, which we can think about. |  |
| Qualcomm | Including EVEX, we can consider the table as the baseline. However, there are more fundamental issues that need to be addressed. So, we would like to propose the following for the future meeting.  **Proposal: RAN2 discussion for analyzing the data collection requirements and solutions for different LCM purposes should be structured. For the data collection for each LCM purpose, companies should provide the following information:**   |  |  |  |  | | --- | --- | --- | --- | | Data collection for LCM purpose | What are the data collection requirements/constraints for the LCM purpose? | Existing Solutions | Why existing solution may work or not? | | Model training | (Below text is just for illustration purpose)  The following constraints may exist for data collection for model training,   * For the development of the model, the data for the model may not be well defined. * Etc…. | SON/MDT | …. | | UE assistance information | …. | | … | …. | |  |  | |  |  | | Model inference |  |  |  | | …… |  |  |  | | ….. |  |  |  |   **NOTE 1: For each LCM purpose, companies should provide their view on the following,**   * **What are the data collection requirements/constraints for the LCM purpose?** * **What are the exiting solutions that companies believe would work for data collection for a given LCM purpose?** * **For each data collection framework for a given LCM purpose, companies should provide why they think it would work or not?** |  |
| Xiaomi | Although payload size is showed in the previous table, but it may not be able to reflect the actual data set each method can deliver. We suggest to show the type of data set and the supported use case of each method. E.g. L3 measurement report may include beam measurement result, which can support positioning and beam management. |  |
| vivo | Support to add the ‘RRC state when performing the data collection’ as a new metric. For instance, for the AI/ML use case of CSI compression, the current non-connected state UE will not perform CSI measurement. |  |
| Spreadtrum | RRC state should be listed in the table. |  |
| CMCC | We are fine to further discuss the aspects proposed by Intel. In addition, we think the RRC state when the UE performs data collection and applicable use cases for each framework can also be considered. |  |

# 3 Conclusion

We will update table based on companies' input.

# 4 References

[1] [R2-2300708](file:///C:\Users\johan\OneDrive\Dokument\3GPP\tsg_ran\WG2_RL2\RAN2\Docs\R2-2300708.zip), Open issues on AI/ML model delivery and data collection in post-meeting email discussion, Apple.