**3GPP T****SG-RAN WG2 Meeting #128 R2-230xxxx**

**Orlando, USA, November 18-22, 2024**

**Agenda item: xxx**

**Source: Interdigital, Nokia**

**Title: [POST127bis][020][AI PHY] Reply LS to SA2/SA5 (InterDigital/Nokia)**

**Document for: Discussion and Decision**

# 1 Introduction

This contribution is aimed at reporting the discussion and results of the following post email discussion:

**[POST127bis][020][AI PHY] Reply LS to SA2/SA5 (InterDigital/Nokia)**

 Intended outcome: Address/discuss SA2 questions from SA2/SA5 LS (if it is sent to RAN2) and possible answers. The discussion is based on RAN2 understanding and previously made agreements. No Tdocs should be submitted to the meeting

 Deadline: Nov. 8th, 10 UTC

Companies providing input to this email discussion are requested to leave contact information below.

|  |  |  |
| --- | --- | --- |
| **Company** | **Name** | **Email Address** |
| Interdigital | Oumer Teyeb | Oumer.teyeb@interdigital.com |
| Nokia | Gyorgy Wolfner | gyorgy.wolfner@nokia.com  |
| ZTE | Fei Dong | dong.fei@zte.com.cn |
| Qualcomm | Rajeev Kumar | rkum@qti.qualcomm.com  |

# 2 Discussion

In [1],RAN sent an LS to SA groups that included the requirements for data collection for a UE-sided AIML model and question regarding which of the data collection solutions identified by RAN2 can fulfil these requirements.

Specifically, the requirements for the data collection indicated in the LS were:

*RAN has agreed to the following* ***requirements for data collection for UE sided model training for standardized solution (if standardized) (i.e. Option 1b, 2, 3). Option 1a is not precluded.***

* + 1. *The data collected is secured and data integrity and confidentiality for that data is ensured.*
		2. *User data privacy, anonymity and user consent is respected.*
		3. *The MNO has full control of the standardized data collection transfer process and can manage data transfer to the server for UE-side data collection, without the need of SLA for this purpose. This includes initiating, terminating, and fully managing data transfer.*
		4. *MNO has full visibility for standardized data.*
		5. *The design is futureproof and extendable.*

*FFS/study if and how to handle non-standardized data (i.e. partial visibility).*

*FFS controllability on data collection*

*Standardized Solutions should follow the principle of aiming to minimize air interface overhead and impact to NW operation*

## 2.1 SA2 LS

In [2],SA2 sent an intermediate response the included some clarification questions. In the following sections, we will address these questions.

### 2.1.1 Controllability of MNO on data transfer

*Q1: Are there any aspects of the UE-data collection controllability, that required NG-RAN involvement? If so, what is the involvement of NG-RAN in UE-data collection controllability, e.g., what aspects of MNO controllability would require NG-RAN involvement and what would such involvement be?*

*As an example of the kind of feedback that is requested, some companies in SA2 understand that initiating (e.g., triggering), terminating collection of UE-side data and controlling data transfer may require NG-RAN involvement, and it is currently not clear what this involvement may be.*

**Rapporteur’s input**

In RAN2-127bis [3], the following agreement was made regarding data collection for model training:

* *Data collection initiation and configuration for data collection is under network control. FFS how the NW determines whether data collection should be initiated (e.g. via UE requests (UE directly or UE server)*

The rapporteur’s understanding is that the NG-RAN is involved in the data collection process, at least for configuring the UE with the required measurements and initiating the data collection.

**A: Do companies agree that the NG-RAN is involved in the data collection procedure, at least in configuring the required measurements and initiating the data collection procedure?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | Yes for configuration;No for initiating data collection procedure | We tend to agree that the NG-RAN involvement includes the RRC configuration related to UE side data collection, which is common understanding in RAN2.For initiating or terminating the data collection procedure , we understand there is no any agreements in RAN2 can deduce such conclusion. In our understanding, the UE side data collection procedure also can be initiated by the CN (e.g. option 2, option 1b) or OAM(e.g. option 3)  |
| Qualcomm | No (with comments) | The question itself is too broad compared to RAN2 agreement. The above RAN2 agreement is only about gNB providing RS configuration and associated ID to the UE, upon UE or UE server request. This procedure for providing RS configuration and associated ID is needed for data collection and applies to all UE side data collection solutions including solution 1a. However, there is no technical reason that it should be part of the data collection configuration and initiation of UE side data collection.As agreed in the RAN2 agreement quoted above, the initiation of UE side data collection is up to the UE or the UE side server.Furthermore, note that UE side can perform training data collection even without training RS configuration and associated IDs.  |
|  |  |  |

If the answer to A is positive, then the rapporteur proposes the following response to Q1 from LS:

*RAN2 confirms that the NG-RAN is involved in the data collection process, and this includes at least providing the UE with the required measurement configurations and initiating the data collection.*

**B: Do companies agree to the proposed response above to Q1 from SA2?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | No | As above comments, we suggest to answer the question as below on top of rapporteur’s suggestion For example:*RAN2 confirms that the NG-RAN is involved in the data collection process, and this includes at least providing the UE with the data collection related configurations.* |
| Qualcomm  | No | RAN2 agreement was about gNB configuring UE with “associated ID” and “RS Configuration for training”. The triggers for data collection and reporting cannot be determined by the network as there are internal UE conditions that determine when the data needs to be collected and reported.Therefore, we suggest modifying the above sentence as:*RAN2 confirms that the NG-RAN should support a procedure for providing RS configuration for training and associated ID to facilitate training data collection (upon UE or UE server request based on RAN2 agreement). The triggers for data collection and reporting cannot be determined by the network as there are internal UE conditions that determine when data needs to be collected and reported.*   |
|  |  |  |

*Q2: Furthermore, with regards to “initiating, terminating and fully managing data transfer” some companies in SA2 believe that further clarification is required, on a per use case basis, on where (which entities) and under what conditions, should controllability be performed, e.g., in NG-RAN, a NF, OAM, an MNO controlled AF, a 3rd party AF, a UE)?*

**Rapporteur’s input**

The rapporteur’s understanding is that the gNB is involved in the UE side data collection for the beam management and CSI prediction/compression use cases, while the LMF is involved for the positioning use cases. This does not mean other entities will not be involved at all in the controlling/enabling the data collection. However, the involvement of other entities outside the RAN is not within the scope of RAN2.

**C: Do companies agree that the gNB is involved in controlling the data collection for the beam management and CSI use cases, while the LMF is involved in controlling the data collection for the positioning use cases?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | No | In our understanding, what we discussed before and having RAN2 agreements is just about the controllability for each option not from use case perspective. We do not think this question can be answered for now from RAN2 perspective. |
| Qualcomm | No | For beam management, upon UE or UE server request, the gNB determines when and what RS configuration and associated IDs for training. Other aspects of UE side data collection are configurable by the gNB. For positioning enhancements, upon the UE or UE server request, the LMF determines when and what PRS configuration and network side addition conditions to provide for training. Other aspects of UE side data collection are not configurable by the LMF. For CSI prediction/feedback, RAN2 should wait for RAN1 discussions.  |
|  |  |  |

If the answer to C is positive, then the rapporteur proposes the following response to Q2 from the LS:

*For the beam management and CSI prediction/compression use cases, at least the gNB is involved in the control of the data collection. For the positioning use cases, at least the LMF is involved in the control of the data collection.*

**D: Do companies agree to the proposed response above to Q2 from SA2?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | No | No answer from RAN2 can be provided for now |
| Qualcomm  | No (suggest modification) | *For the beam management, the gNB should support a procedure for providing the RS configuration and the associated IDs for training, based on UE or UE server request. For the positioning use cases, the LMF should support a procedure for providing the PRS configuration and the network side additional conditions for training, based on UE or UE server request. For CSI prediction/compression use cases, the gNB support for providing RS Configuration and associated ID is still under RAN1 discussion. The triggers for data collection and reporting cannot be determined by the network as there are internal UE conditions that determine when data needs to be collected and reported.*  |
|  |  |  |

*Q3: Furthermore, some companies in SA2 wondered whether full controllability would have any impact on UE normal operation. If so, what impact is expected from RAN2 perspective to enable UE-side Data Collection?*

**Rapporteur’s input**

For collecting data for the training of a network side model, RAN2 is already discussing the impact on UE’s performance and how to minimize that. For example, the usage of lower priority SRB for sending the collected data has already been agreed to ensure that data reporting will not delay other important control plane message.

Our understanding is that there will be some impact to the UE’s operation/performance due to the data collection/reporting for UE side model training. The level of impact on UE’s normal operation, as well as other factors such as the impact on the NW, air interface load, specification impact, etc., may also be considered to down select among the identified solutions that enable MNO controllability. Also, considerations must be made in the design of the final data collection solution to ensure the impact on the UE’s performance/operation are minimized.

However, the question from SA2 seems to be on the impact of the full controllability aspect on UE’s operation, rather than the general aspect of UEs performing the data collection and reporting. During the RAN2 discussions so far, no impact on UE’s normal operation due to the full controllability of the data collection process has been identified.

**E: Do companies agree that no direct impact on UE’s normal operation due to the full controllability of the data collection process has been identified by RAN2?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | See comments | What is the UE’s normal operation？What kind of UE behavior can be called as normal operation, We are confused about such definition from SA. |
| Qualcomm | No | This was not discussed in RAN2. As the data collection procedure may place a huge burden on the UE, only the UE side may determine when/what is the appropriate time/conditions for the collection/reporting of the training data. The UE side may consider the UE hardware version, software version, load, power, memory, and other factors in account when initiating the collection/ reporting of the training data.We also agree that in solution 1b/2/3, the network side can also consider when/what is the appropriate time/conditions for the reporting of the training data. In option 1a, it is up to UE implementation. Coming to the response to the SA2 question, if the full controllability means that network decides when does the UE collect and report the training data (in solution 1b/2/3), that may impact the UE normal operations. The RAN doesn’t have enough information to mitigate the impact.As discussed, previously, the network may determine when it want to provide RS configuration for training and associated IDs for training, but it is up to UE whether UE can and wants to collect training data and triggers for UE side data collection.  |
|  |  |  |

If the answer to E is positive, then the rapporteur proposes the following response to Q3 from the LS:

*RAN2 has not identified any impact on UE normal operation due to the full controllability.*

**F: Do companies agree to the proposed response above to Q3 from SA2?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | No | We need to ask SA what is UE normal behaviour, and what kind of UE behaviour can be called as normal operation... |
| Qualcomm | No (suggest modification) | *There may be impact on UE normal operation due to the full controllability. Only the UE can determine appropriate time/conditions for UE-side training data collection/reporting. In solution 1b/2/3, the UE report the collected data based on network-provided configurations and UE-determined time/conditions for UE-side training data reporting.*  |
|  |  |  |

*Q4: Some companies in SA2 understands that standardized data content refers only to data reflecting results of measurements performed by the UE according to network measurement configuration. SA2 would kindly asks RAN2 to confirm this understanding.*

**Rapporteur’s input**

Most of the standardized data to be collected is expected to be according to measurement configuration provided by the network. However, there may be elements in the report that are not based on measurement configuration provided by the network. For example, timestamp information is indicated to be one of the information elements to be collected/reported for the beam management case.

What is meant by standardized data is that the format and the meaning of the data will be known by the network (e.g., the type of information that is contains, the size/type of the data, etc.,).

Thus, the rapporteur proposes the following response to Q4 from the LS:

*Most of the collected/reported standardized data will be according to the measurement configuration provided by the network. However, there could be information elements (e.g., timestamps) in the collected/reported data that may not be acquired based on the measurement configuration. Thus, standardized data can be defined without necessarily tying it to measurement configuration and it refers to data whose format will be explicitly defined in 3GPP specifications, and the network will be able to understand the content/meaning of the data based on that.*

**G: Do companies agree to the proposed response above to Q4 from SA2?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | Yes | This could be RAN2 common understanding. |
| Qualcomm | No (suggest modification) | *A part of the collected/reported standardized data will be according to the measurement configuration provided by the network. However, there could be additional information elements (e.g., timestamps) in the collected/reported data that may not be standardized. The standardized data will be explicitly defined in RAN1/RAN2 standard specifications.*  |
|  |  |  |

### 2.1.2 Roaming support

*Q5: Does RAN2 expect data to be collected from UEs that are roaming?*

**Rapporteur’s input**

Roaming considerations are in general outside the scope of RAN2.

The UE is not operating autonomously (i.e., the network is the one that configures the measurements, controls the data collection/transfer). Thus, it is up to the network to enable/disable the data collection operation when the UE is roaming.

Any further aspects of roaming considerations are in general outside the scope of RAN2.

The rapporteur proposes the following response to Q5 from the LS:

*The UE is not operating autonomously (i.e., the network configures the required measurements and controls the data collection/transfer). Thus, it is up to the network to enable/disable the data collection operation when the UE is roaming. Any further aspects of roaming considerations are in general outside the scope of RAN2.*

**H: Do companies agree to the proposed response above to Q5 from SA2?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | No | RAN2 does not touch this discussion, and no RAN2 any agreement can reflect above response. We would like to answer this question simply:*No conclusion about roaming is reached in RAN2.* |
| Qualcomm | No | Whether roaming can be supported or not is under SA2/SA3 scope. For example, RAN is not aware about HPLMN and VPLMN, and it should be decided by SA2/SA3 whether training data collection can be supported across VPLMN and HPLMN based on bilateral agreements and local regulatory. Therefore, suggest removing the first two paragraph and rewording, as below*Roaming considerations are outside the scope of RAN2.* |
|  |  |  |

### 2.1.3 Visibility

*Q6: SA2 would like confirmation from RAN2, whether it is sufficient that the data content be standardized and MNO knowing which data type is collected, and the MNO knowing the actual data content to be considered as the MNO having full visibility of such data or should further conditions be required in order to declare that the MNO has full visibility of such data, e.g. whether MNO need to verify the match between the data transferred and the data collected ?*

**Rapporteur’s input**

As stated in the LS sent from RAN and also discussed above in relation to Q4, visibility of data content only signifies that the MNO will be able to be aware of, access, and comprehend the content of the collected/reported data without the need of SLA.

Also, it is not clear on how the MNO will be able to verify the match between the data transferred and the collected data. The data is collected at the UE mainly because it is information (e.g., such as DL signal levels) that only the UE can measure. If the concern is about the quality/accuracy of the collected data, RAN4 discussions are being (will be) made concerning that and UEs performing data collection must comply with any requirements that will be set based on that discussion (e.g., like current requirements for RRM measurement reporting).

Thus, there are no further requirement for the MNO to verify the match between data transferred and data collected.

The rapporteur proposes the following response to Q6 from the LS:

*As stated in the LS sent from RAN, visibility of data content only signifies that the MNO will be able to be aware of, access, and comprehend the content of the collected/reported data without the need of SLA. If the concern is about the quality/accuracy of the collected data, UEs performing data collection must comply with any requirements that will be set in RAN4 (e.g., like current requirements for RRM measurement reporting). Thus, there are no further requirement for the MNO to verify the match between data transferred and data collected.*

**I: Do companies agree to the proposed response above to Q6 from SA2?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | No | In our understanding, the intention of full visibility is for MNO to check whether the data transferred to the UE server is matched to the data collected based on collection configuration to avoid the potential privacy leakage.In this sense, we think there is further requirement for the MNO to verify the match between the data transferred and the data collected. |
| Qualcomm | No (suggest rewording) | This is about UE-side data collection. Therefore, the network does not need to be check the quality. Suggest rewording as: *As stated in the LS sent from RAN, visibility of data content only signifies that the MNO will be able to be aware of, access, and comprehend the content of the collected/reported data without the need of SLA. There are no further requirement for the MNO to verify the match between data transferred and data collected.*  |
|  |  |  |

## 2.4 SA5 LS

In [4]**,** SA5 sent a reply LS to the RAN LS on AIML data collection (RP-242389) including two questions.

*Q8: Is the “Server for data collection for UE-side model training” controlled by operators?*

**Rapporteur’s input**

The requirement is for the data collection to be fully controlled by the MNO without the need for the SLA, and it is not about controlling the server. As long as the MNO can do that, a solution can be considered feasible by SA5 even if the server for data collection is located outside the MNO’s domain.

Thus, the rapporteur proposes the following response to Q8:

*The controllability requirement is referring to the controlling of the data collection/transfer process, and it is not concerned about the control of the server for data collection.*

**J: Do companies agree to the proposed response to Q8 above (Q4.1, part 1 from SA5)?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | See comments | I guess SA5 is actually to ask the ownership of the server for data collection for UE side model training that we have discussed in RAN2 before but there is no any conclusion is made. We can answer the question directly:RAN2 does not reach the consensus about the controlling of server for data collection.  |
| Qualcomm | Yes (with comment) | From the RAN2 perspective, *the controllability requirement is referring to the controlling of the data collection/transfer process. Whether the the “Server for data collection for UE-side model training” controlled by operators is outside RAN2 discussion.* *Suggest modification:*From the RAN2 perspective, *the controllability requirement is referring to the controlling of the data collection/transfer process. Whether the “Server for data collection for UE-side model training” is controlled by operators or not, is outside RAN2 discussion/scope.*  |
|  |  |  |

*Q9: What standardized data is to be collected?*

**Rapporteur’s input**

As indicated in the LS from RAN, RAN1 has provided initial information about the data to be collected for the different use cases (R1-2310681). However, this is initial information, and it is likely that further updates/additions will be made as the work/study item progresses. For example, RAN1 has made further agreements regarding the data to be collected for the positioning use case in RAN1#116b [5]. However, it is reasonable to assume that the amount of data to be collected (per sample) will be similar as the provided in R1-2310681.

Thus, the rapporteur proposes the following response to Q9:

*SA5 can refer to R1-2310681 for the content of standardized data to be collected for the different AIML use cases. Further updates/addition to this are likely to be made as the work/study item progresses. However, SA5 can assume the size of the data to be collected (per sample) will be similar as the one provided in R1-2310681.*

**K: Do companies agree to the proposed response to Q9 above (Q4.1, part 2 from SA5)?**

|  |  |  |
| --- | --- | --- |
| **Company** | **Yes/No** | **Comments** |
| ZTE | Yes |  |
| Qualcomm | Yes (with comments) | Agree with Rapporteur. Suggested modification:*SA5 can refer to R1-2310681 for the content of standardized data to be collected for the different AIML use cases. There can be additional contents that can be collected at the UE for UE side model training, as mentioned in R1-2310681. Further updates/addition to this are likely to be made as the work/study item progresses. However, SA5 can assume the size of the data to be collected (per sample) will be similar as the one provided in R1-2310681.*  |
|  |  |  |

# 3 Conclusion

To be added...

# 4 Reference

1. RP-242389 /S2-2409600, LS on AIML data collection, RAN #105, September 2024
2. S2-2411191, Reply LS on AIML data collection, SA2 #165, October 2024
3. RAN2 #127-bis, Chairman Notes, October 2024
4. S5-246299, Reply LS on AIML data collection, SA5 #157, October 2024
5. RAN1 #116-bis, Chairman Notes, April 2024