3GPP TSG-RAN WG2 Meeting #126 R2-240xxxx

Fukuoka, Japan, May 20th -24th, 2024

Agenda Item: 8.1.4

Source: Mediatek Inc.

Title: Report of [AT126][020][AI/ML PHY] UE side data collections

Document for: Discussion, Decision

# Introduction

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

* [AT126][020][AI/ML PHY] UE side data collections (Mediatek)

 Intended outcome: Agreeable table for UE side data collection and clarification of visibility, levels of visibility, and standardized vs. non-standardized

 Deadline: 05-24-24

# Discussion

During the online discussion, the different levels of visibility on the data content were discussed, i.e., full visibility, partial visibility, and no visibility.

To achieve different levels of visibility, we need to assess whether the data content is standardized, non-standardized, or partially standardized. An instance of partially standardized data content could be an Information Element (IE) with an undefined value where the fields are visible but the exact values remain undisclosed. Based on the variations in data content visibility and standardization levels, the following possibilities have been identified:

1. Full visibility for standardized data content.
2. Full visibility for non-standardized data content as per the SLA.
3. Partial visibility for partially standardized data content.
4. Partial visibility for non-standardized data content as per the SLA.
5. No visibility for non-standardized data content.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Full Visibility | Partial Visibility | No Visibility |
| Standardized data content | 1. Yes | NA | NA |
| Partial-standardized data content | NA | 3. Yes | NA |
| Non-standardized data content | 2. Can be visible based on SLA or business contract providing data description | 4. Can be visible based on SLA or business contract providing data description | 5. Yes |

Those possibilities are applicable to solution 1b, 2 and 3.

**Proposal 1: For solution 1/2/3, the following options are identified to realize the different levels of data content visibility:**

1. **Full visibility for standardized data content.**
2. **Full visibility for non-standardized data content as per the SLA.**
3. **Partial visibility for partially standardized data content.**
4. **Partial visibility for non-standardized data content as per the SLA.**
5. **No visibility for non-standardized data content.**

**Proposal 2:** **The MNO has full visibility of the data content either through standardized data format or via SLA/business contract for partial-standardized/non-standardized data for solution 1b, 2 and 3.**

For solution 1b, as outlined by Ericsson and Qualcomm, varying degrees of data content visibility can be established through SLA. The visibility, whether full, partial, or none, is not affected by whether the data content is standardized or not since it is stipulated by the SLA and transmitted via a UP tunnel. Thus, both standardized and non-standardized data can be configured to have full, partial, or no visibility, in accordance with the terms set forth in the SLA.

**Proposal 3: RAN2 assumes that** **different levels of visibility on the data content can be achieved via SLA defined by SA2 in solution 1b.**

**Proposal 4: RAN2 assumes that different levels of visibility on the data content can be achieved via SLA or business contract in solution 2/3 for partial-standardized/non-standardized data content.**

**Proposal 5: As a starting point, RAN2 assumes that 'visibility' of data content signifies the capability of the MNO to, at least, be aware of, access, and comprehend the data during transfer. The scope does not exclude additional requisites, such as the ability to modify the collected data. FFS on the meaning of modify. NOTE: It is an assumption for RAN2 study purpose. The definition of “visibility” and the category of visibility should be discussed and defined in SA1.**

**Proposal 6: RAN2 endorse Table 1 to capture the characteristics of different options for UE-side training data collection as the starting point for future discussion.**

**Proposal 7: Capture the privacy concerns from different stakeholders as informative annexes in the TR.**

Table 1 Characteristics of different options for training data collection for UE-side models

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Aspects | 1a) OTT (3GPP Transparent) | 1b) The server for training data collection for UE-side models (3GPP non-transparent) | 2. Transfer via Core Network | 3. Transfer via OAM |
|  |  |  |  |  |
|  |  |  |  |  |
| UP/CP tunnel | UP tunnel (Note: data collection may be charged as normal traffic.) | UP tunnel (Note: data collection may be charged as normal traffic.) | CP tunnel (provided the data volume remains within the NAS signalling capacity)FFS: UP tunnel | CP tunnel (provided the data volume remains within the RRC signalling capacity)FFS: UP tunnel |
| Data Transfer Path | NA | UE->gNB->CN (FFS on NF)/AF->Server for data collection for UE-side model training/OTT server | UE->gNB->CN (FFS on the NF)-> Server for data collection for UE-side model training/OTT server | UE->gNB->OAM-> Server for data collection for UE-side model training/OTT server |
| Protocol layer for data transfer | Application layer | Application layer | NAS layer for CP tunnelFFS: the protocol layer for UP tunnel | RRC layer for CP tunnelFFS: the protocol layer for UP tunnel UP tunnel |
| Controllability of MNO on data transfer | No specific controllability | Has controllabilityFFS: level of controllability | Full controllability (Note 1) | Full controllability (Note 1) |
| Control granularity by NW | NA, the OTT server can directly request data from the UE. | Example: per PDU sessions based on SLA | NAS procedure | RRC procedure |
| Visibility of data content in MNO | No visibility | Full visibility (Note 2, Note 3, Note 4) | Full visibility (Note 2, Note 3, Note 4) | Full visibility (Note 2, Note3, Note 4) |
| Data format  | Out of 3GPP scope | Standardized and non-Standardized | StandardizedFFS: non-standardized | StandardizedFFS: non-standardized |
| Involved WGs | No, out of 3GPP scope | SA2, RAN2 | SA2, RAN2 | SA5, SA2, RAN2 |
| * Note 1: Full controllability: The MNO has the capability to manage data transfer to the server for UE-side data collection. This includes initiating, terminating, and fully managing the volume of data. (Subject to refinement and modification)
* Note 2: Visibility of data content signifies the capability of the MNO to, at least, be aware of, access, and comprehend the data during transfer. (Subject to refinement and modification, the scope does not exclude additional requisites, such as the ability to modify the collected data.)
* Note 3: The MNO has full visibility of the data content either through standardized data format or via SLA for non-standardized data in solution 1b, 2 and 3.
* Note 4: Different levels of visibility on the data content can be achieved via SLA defined by SA2 in solution 1b or SLA/business contract in solution 2/3..
* The following options are identified to realize the different levels of data content visibility if different levels of data content visibility to MNO are considered:
1. Full visibility for standardized data content.
2. Full visibility for non-standardized data content as per the SLA.
3. Partial visibility for partially standardized data content.
4. Partial visibility for non-standardized data content as per the SLA.
5. No visibility for non-standardized data content.
 |