3GPP TSG-RAN WG2 Meeting #126 R2-24xxxxx

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

Agenda Item: 8.1.4

Source: Mediatek Inc.

Title: Report of [POST125bis][020][AI/ML PHY] UE side data collection

Document for: Discussion, Decision

# Introduction

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

* [POST125bis][020][AI/ML PHY] UE side data collection (Mediatek)

Intended outcome: Discuss new table capturing solution details and discussion fon control and visibility, privacy.

Deadline: two weeks (Deadline May 3rd, 10:00 UTC)

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

|  |  |  |
| --- | --- | --- |
| **Company** | **Name** | **Email Address** |
| NEC | Xuelong Wang | xuelong.wang@emea.nec.com |
| Apple | Peng Cheng | Pcheng24@apple.com |
| BT | Salva Diaz | salva.diazsendra@bt.com |
| Nokia | Jerediah Fevold | jerediah.fevold@nokia.com |
| Ericsson | Marco Belleschi | marco.belleschi@ericsson.com |
| Huawei, HiSilicon | Jun Chen | jun.chen@huawei.com |
| OPPO | Jiangsheng Fan | fanjiangsheng@oppo.com |
| Mediatek | Yuanyuan Zhang | Yuany.zhang@mediatek.com |
| vivo | Boubacar Kimba | kimba@vivo.com |

# 2 Discussion

The following proposals were identified during the Rel-18 SI on AI/ML for NR Air Interface:

|  |  |
| --- | --- |
| 1. UE collects and directly transfers training data to the Over-The-Top (OTT) server;   1a) OTT (3GPP transparent)  1b) OTT (non-3GPP transparent)   1. UE collects training data and transfers it to Core Network. Core Network transfers the training data to the OTT server. 2. UE collects training data and transfers it to OAM. OAM transfers the needed data to the OTT server.   RAN2 did not study or analyse these proposals and did not agree to requirements or recommendations. | |
| 1a) | 1b) |
| 2. Data collection via CN | 3. Data collection via OAM |

According to the consensus made during the online discussion in RAN2#125bis meeting,

=> need to better define what is control of data collection in MNO and visibility of data content in MNO.

=> understanding is that OTT is outside of MNO

We will clarify these methods in the following aspects:

* Inside/outside MNO’s network
* Termination Entity
* Controllability of data collection in MNO
* Visibility of data content in MNO
* Protocol layer for data transfer
* Privacy concerns

## 2.1 Terminology

According to the common understanding made in RAN2 that OTT is outside of MNO, it is inappropriate to refer to the term OTT in the context of solutions 1b, 2, and 3. To avoid confusion and enhance clarity, we propose replacing the term ' OTT server' with 'server for UE-side data collection' in these solutions. The term "UE-side data collection" refers to the data collection intended for use in UE-side model training. Please note that the term 'training' is omitted from the terminology to prevent confusion or debate regarding whether the locations for data storage and model training are identical or separate.

**Q1:** **Do companies agree to replace the term ‘OTT server’ with ‘server for UE-side data collection’ in the definitions/descriptions of different solutions?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | Yes for solution 1b/2/3 with comments:  As this is new terminology created by RAN2, we are still confused with whether it is:   1. a server owned by UE vendor, or 2. a server owned by MNO, or 3. a neutral server   Since **the privacy requirements for the 3 possible options are different**, we think it needs separate discussion. One may argue it is out of RAN2 scope, but please note that RAN1 has agreed and discussed extensively on different model transfer solutions (y and z1-z5) depending on whether training is UE-sided, NW-sided or neutral site.    Thus, similarly, we think it is one important aspect RAN2 should discuss. Meanwhile, it is also part of “solution details” which are explicitly included as scope of this email discussion.  [Rapp1] This definition is very general and doesn’t imply any ownership of the server. The ownership is intended to be clarified whether it’s inside or outside MNO’s network. |
| Nokia | Yes, but maybe to be even clearer, we could use the term “server for training data collection for UE-side models”. |
| Ericsson | Yes, however the terminology “OTT server” is used to define a node outside the MNO premises. Hence (similar to Huawei comment below) for solution 1a) we should refer the server for UE-side data collection outside the MNO. Whereas for the solutions 1b), 2, 3 the server for UE-side data collection can be inside or outside the MNO premises.  [Rapp1] Yes, with the change of the terminology, the general description of solution 1a, 1b, 2, and 3 needs to be updated accordingly.  OK with Nokia clarification as well “server for training data collection for UE-side models outside the MNO”. |
| Huawei, HiSilicon | Yes with comments.  In our understanding, if we keep the original text in TR 38.843, the termination is OTT server. If we introduce another terminology for the server, we would like the email rapporteur to confirm whether our understanding is correct or not:  **If this "server for UE-side data collection" is inside MNO, the server still needs to transfer the collected data to UE server outside MNO (for training purpose).**  [Rapp1] As I emphasized initially, my goal was to avoid any ambiguity or debate concerning whether the sites for data storage and model training are the same or separate. This issue seems to be a subsequent inquiry. In my view, if we're discussing a 'server for UE-side data collection,' inside MNO, it should be understood that the data gathered is retained within the MNO's network for training purposes as well. This arrangement would typically involve the UE vendor renting server space from the MNO, or the MNO buying a server from the UE vendor and setting it up within their own network. |
| OPPO | For solution 1a, we still can use OTT server, while for the other three solutions, i.e. 1b, 2,3, we can use ‘server for UE-sided data collection’. |
| Mediatek | Yes. The terminology proposed by Nokia is more precise but is somewhat lengthy and potentially cumbersome for ongoing discussions. Perhaps we could use the abbreviated term 'server for UE-side data collection' for the time being. It would be beneficial if the TR rapporteur could furnish a definition for this term, aligning with Nokia's suggestion, to ensure clear understanding |
| vivo | Yes for solutions 1b, 2, and 3. |

## 2.2 Inside/outside MNO’s network

One point that requires clarification is the location of the server for UE-side data collection relative to the MNO’s network. In solution 1a, the UE directly communicates with the server for UE-side model collection, typically via the application layer, without any interaction with the MNO's network infrastructure. Consequently, in this scenario, the server for UE-side data collection operates outside of the MNO's network, functioning as an OTT server. However, for solutions 1b, 2, and 3, the server for UE-side data collection has the flexibility to be positioned either inside or outside the MNO's network. It is noted that in solution 2, the server for UE-side data collection can be a NF in CN.

Reflecting on the feedback received, there's a need to clarify the intention and meaning on whether the server is located inside or outside the MNO's network. The core of this clarification is to determine server ownership, as this directly influences who is responsible for adhering to privacy laws, regulations, and so forth. If the server resides inside the MNO's network, it is considered to be owned by the MNO. Conversely, if it is located outside the MNO's network, ownership falls to the UE vendor or another external party, but not the MNO. So, I add one more question to align companies’ understanding.

Q2.0: Do companies agree that a server located within the MNO's network is deemed to be MNO-owned, whereas one located outside is not under MNO ownership?

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| Mediatek | Yes |
| vivo | Yes |
|  |  |
|  |  |
|  |  |
|  |  |

|  |  |
| --- | --- |
| **Solutions** | **Inside/outside MNO’s network** |
| **1a** | OTT server, i.e., outside MNO’s network |
| **1b** | Inside/outside MNO’s network |
| **2** | Inside/outside MNO’s network  Note: The server for UE-side data collection can be a NF in CN. |
| **3** | Inside/outside MNO’s network |

**Q2.1:** **Do companies agree that for solution 1a the server for UE-side data collection is outside of MNO’s network and is therefore classified as an OTT server?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | Yes, as captured in Chair Note. |
| BT | Yes but for us, solution 1a is totally outside 3GPP scope. Consequently, it should be taken away from the WI phase. No need to discuss option 1a. |
| Nokia | Yes. We would also like to know if the following figure, an excerpt from TS26.531, Figure 4.2-1 is the intended representation of solution 1a.    Additionally, the following note is provided in the same specification: “NOTE 5: Interactions at reference point R8 are beyond the scope of 3GPP standardisation.” |
| Ericsson | Yes |
| Huawei, HiSilicon | Yes with comments.  **The wording could be improved, i.e. suggest to use outside-MNO server.** |
| OPPO | Yes |
| vivo | Yes |

**Q2.2:** **Do companies agree that for solutions 1b, 2 and 3, the server for UE-side data collection can be either inside or outside MNO’s network?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Solution 2, Yes. But,  My understanding is that solution 1b) should still be sub-option of the solution 1 group (i.e. UE directly transfers data to server), so the difference from the solution 1a) is the server is “inside” MNO’s network and solution 1b is that the server is “outside” MNO’s network. We propose to preclude “outside” from 1b).  In addition, for solution-3, it is a bit unclear if the UE-side data collection should be inside or outside MNO’s network. |
| Apple | No for solution 1b, Yes for solution 2 with comments, not sure for solution 3.  On solution 1b, we are not sure how the server can be inside MNO’s network. As we commented in Q1, we think two solutions can be:   * a) UE vendor can deploy its owned server in MNO network or   + In this case, which Network entity UE vendor can deploy its server (e.g. in CN or RAN)? * b) MNO deploy some servers (i.e. owned by MNO) which are rent to UE vendor for data collection?   Note that in case of understanding a), it may have 3GPP impact and cross-WG impact. Thus, RAN2 should make it clear. Before this issue is concluded, we think “inside” should be precluded from solution 1b).  For solution 2, we agree it can be inside or outside MNO network. Meanwhile, we appreciate that Rapporteur provides example of “a NF”, but we are not whether it means “a NF” performs model training on behalf of UE vendor or it is just an intermedium entity for further forwarding dataset? Please clarify.  [Rapp1] As I emphasized initially, my goal was to avoid any ambiguity or debate concerning whether the sites for data storage and model training are the same or separate. This issue seems to be a subsequent inquiry. In my view, it’s just an intermedium entity for further data forwarding.  For solution 3, we agree with NEC it is not clear. We noticed Rapporteur doesn’t provide example. We are wondering whether anyone can give a concrete example that the final UE server is within MNO network (e.g., if it is an entity in RAN, is it within DU or CU)?  [Rapp1] Please check my example provided below in Mediatek’s comment |
| BT | No.  From our understanding, the server for UE-side collected data on options 1a, 2 and 3 are inside the MNO and the ownership is for the MNO.  A different discussion is where the algorithms for training, management and inference are allocated but we consider this is not asked here. |
| Jerediah [Nokia] | Maybe, but we first need to clarify what “outside” the network means, and whether that server could have, e.g., a VPN or N3IWF interface to the NW to communicate with NFs that are involved in DC. And it should be clarified whether there is any other difference in the interface between the server for UE-side data collection whether it is located inside or outside of the network. |
| Ericsson | Yes. For all solutions (except 1a), depending on the deployment and specific SLA agreement, the server for UE-side data collection can be inside or outside the MNO network. In case the server for UE-side data collection is outside the MNO, then it should be discussed for all the cases (except 1a)), what is the involvement of 3GPP network nodes. |
| Huawei, HiSilicon | No.  Firstly, we have the same comments as for Q1, and we would like to confirm our understanding as below.  **If this "server for UE-side data collection" is inside MNO, the server still needs to transfer the collected data to UE server outside MNO (for training purpose).**  [Rapp1] Regardless of the server's location, whether within or outside the MNO's network, the interface connecting the server to entities within the MNO may either be standardized or non-standardized interface. I believe that assessing the implications on specifications solely from a RAN2 perspective is challenging. My suggestion is that we delve into the solutions' details as thoroughly as possible from the RAN2 standpoint and then engage other WGs to assess the impacts on specifications. We should pose the question once we have a clearer understanding of each solution's specifics.  Secondly, we suggest to clarify the impacts if UE server is deployed inside MNO, and also to clarify differences between inside and outside MNO from a standard point of view.  Thirdly, we suggest to clarify the following: if the training data (from UE side) is transferred to a UE-side server inside MNO, does it mean that MNO can directly use the data for training purpose?  [Rapp1] The data is accessible to MNO, and MNO can use data for training if it has interest. But the MNO’s effort would be huge if MNO takes the training responsibilities for different UE vendors/chipsets with different hardware/firmware requirements. In practice, it may involve the UE vendor renting server space from the MNO, or the MNO buying a server from the UE vendor and setting it up within their own network. (This part can be commented by operators). |
| OPPO | Yes for solution 2 and solution 3, some concern for solution 1b.  For solution 2 and solution 3, OPPO understanding is that usually specified data type will be used to collect UE side data, in this case, both operator and UE can decode the collected data, so it makes sense that the server for UE-side data collection can be either inside or outside MNO’s network as both operator and UE server can use the data without extra offline engineering work.  When it comes to solution 1b, we understand that usually unspecified data type will be used to collect UE side data, in this case, only UE vendor can decode the collected data, so it doesn’t make sense that the server for UE-side data collection is inside MNO’s network as operator can not use the collected data without extra offline engineering work. It makes more sense that for solution 1b, server for UE-side data collection is outside MNO’s network. |
| Mediatek | Yes.  For solution2, the server can be a NF if it is inside MNO’s network. If the server is outside MNO’s network, there may be an interface between the server and CN.  For solution 3. If the server is outside of MNO’s network, there may be an API between the server and OAM. If the server is inside of MNO’s network, it may be similar as TCE, which may be even within OAM domain.  For all solutions, the interface connecting the server to entities of the MNO may either be standardized or non-standardized interface. We should figure out the solutions' details as thoroughly as possible from the RAN2 standpoint and then engage other WGs to assess the impacts on specifications. |
| vivo | The boundary between inside and outside MNO’s network should be further clarified. Besides, inside or outside MNO’s network is up to SA.  From RAN2 perspective, the key characteristic of option 1b/2/3 is that the server for UE-side data collection is known by a specific MNO entity, and the MNO entity is involved in the data exchange between UE and the server. |

## 2.3 Termination Entity

The "(First)termination entity" refers to the entity that receives and stores data transmitted from the UE, **which possesses the authority to oversee the subsequent handling of this data**, such as data cleaning, forwarding, sharing, and analysis, among others, in compliance with privacy policies, security protocols, and any regulatory compliance requirements. The purpose of introducing the term ‘(first) termination entity’ is to emphasize the entity’s role in management and possession of the collected data.

According to the responsibility of the termination entity as define above, the termination entity for solution 1a) is the OTT server. The termination entity for solution 1b is the server for UE-side data collection. The termination entity for solution 2 is the CN. The termination entity for solution 3 is the OAM.

**Q3.1: Do companies agree that for solutions 1a the termination entity of UE-side data collection is the OTT server?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | Yes |
| BT | Yes but as it was mention before, no need to discuss 1a |
| Nokia | Yes. |
| Ericsson | Yes. By the definition given above by the rapporteur, the termination entity is inside the MNO. We propose clarifying that, and use instead the terminology “termination entity inside the MNO”, rather than “first termination point”. This is because as per what captured in the TR, in all the approaches the data are ultimately transferred/terminated to the server for UE-side data collection (previously named OTT server) which could be outside the MNO. So it would be good to clarify that in all the approaches (except 1a), there might be a termination entity inside the MNO, and one outside (i.e. the training entity). Whereas for option 1a), the termination entity is only outside the MNO.  [Rapp1] Tend to agree. It means for solution 1a, there is no termination entity inside the MNO. Let’s check whether there are any other suggestions. |
| Huawei, HiSilicon | Yes |
| OPPO | Yes, better to add the term ‘first’ before termination entity to remove ambiguity. So does for the following questions. |
| Mediatek | Yes |
| vivo | Yes |

**Q3.2: Do companies agree that for solutions 1b the termination entity is the server for UE-side data collection?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | Yes |
| BT | Q2.2 indicates “for solutions 1b, 2 and 3, the server for UE-side data collection can be either inside or outside MNO’s network”.  We do not consider “outside” the MNO as an option for that reason, we propose to clarify the sentence:  Do companies agree that for solution 1a, the termination entity of UE-side data collection is inside the MNO?  Using our proposed sentence, the answer is yes. |
| Nokia | Yes. We would also like to know if the following figure, an excerpt from TS26.531, Figure 4.2-1 is the intended representation of solution 1b, wherein the first termination entity is actually in the CN. Otherwise, we would like to discuss what is the difference between what is depicted in the figure vs. solution 1b as depicted in the diagram (shown in the introduction of Section 2, top-right) for solution 1b provided in this Email Discussion. |
| Ericsson | Yes, but the terminology “termination entity” might be a bit ambiguous (as per the previous comment). Propose to clarify “termination entity inside the MNO is the server for UE-side data collection”.  As commented above, in all the approaches there might be always a “second” termination entity (former OTT server in charge of training) which is outside the MNO. Hence it should be clarified that the “first” termination entity discussed in this section is inside the MNO. |
| Huawei, HiSilicon | Yes with comments.  For Solution 1b, the terminology “termination entity” is a bit confusing. As we asked in Q1, whether there is one UE server or more than one UE server should be clarified. |
| OPPO | Yes |
| Mediatek | Yes |
| vivo | Yes |

**Q3.3: Do companies agree that for solutions 2 the termination entity of UE-side data collection is the CN?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | No.  According to the nice figure provided by Rapporteur, only the first termination point is CN in Solution 2 (i.e. final termination entity is still “server for UE-side data collection”). Thus, we suggest below rewording:  **for solutions 2 the first termination entity of UE-side data collection is the CN** |
| BT | For clarification, we propose the following sentence  Do companies agree that for solutions 2 the termination entity of UE-side data collection is inside the CN?  Yes. That was the proposal done on [R2-2403492](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403492.zip) |
| Nokia | Yes, however it is unclear which CN entity is the termination entity. We would also clarify that the termination in the CN is not the UPF. |
| Ericsson | Yes, but the terminology “termination entity” might be a bit ambiguous (as per the previous comment). Propose to clarify “termination entity inside the MNO is a CN node”.  In all the approaches there might be always a “second” termination entity (former OTT server in charge of training) which is outside the MNO. Hence it should be clarified that the “first” termination entity discussed in this section is inside the MNO. |
| Huawei, HiSilicon | Yes with comments.  **We agreed with Apple that first termination entity should be used in order to avoid confusions.** |
| OPPO | Yes |
| Mediatek | Yes. Tend to agree with Ericsson to clarify that the termination entity inside the MNO is a CN node. |
| vivo | Yes |

**Q3.4: Do companies agree that for solutions 3 the termination entity of UE-side data collection is the OAM?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | No.  Similar to Q3.3, we suggest below rewording:  **for solutions 3 the first termination entity of UE-side data collection is the OAM** |
| BT | Similar to Q3.3, we propose:  Do companies agree that for solutions 3 the termination entity of UE-side data collection is inside the OAM?  Yes. The UE-side data collected is still under MNO control. |
| Nokia | Yes. |
| Ericsson | Yes, but the terminology “termination entity” might be a bit ambiguous (as per the previous comment). Propose to clarify “termination entity inside the MNO is the OAM”.  In all the approaches there might be always a “second” termination entity (former OTT server in charge of training) which is outside the MNO. Hence it should be clarified that the “first” termination entity discussed in this section is inside the MNO. |
| Huawei, HiSilicon | Yes with comments.  **We agreed with Apple that first termination entity should be used in order to avoid confusions.** |
| OPPO | Yes |
| vivo | Yes |

## 2.4 Controllability for transfer of the collected data in MNO

To begin, we focus on the aspect of data transfer to address controllability.  Once these aspects are clearly understood, we can delve into more intricate details, such as how the MNO might configure measurement and reporting for the UE, among other considerations.

We define the controllability of data transfer for UE-side data collection within an MNO in the following dimensions:

* to and from the server for UE-side data collection
* The specific entity within the MNO to control the data transfer to and from the server for UE-side data collection.
* The protocols and methods utilized by the MNO to control the data transfer to and from the server for UE-side data collection.

’Companies are invited to discussion on whether it is appropriate to start discussions on the controllability of data transfer, as outlined in the preceding four bullet points. Additionally, companies are requested to provide any definitions of controllability that may encompass critical aspects not already covered.

**Q4.1** **Do companies agree to start controllability discussion on data transfer as defined with the above four bullets? Please provide/clarify the definition on controllability if there are any important aspects missing.**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes, but it would be important to clarify which entity within the MNO that should be responsible for the control of the data transfer for each solution if any. |
| Apple | Yes with comments:  We understand the 1st dimension is covered by the 4th one. So, suggest to remove 1st dimension. |
| BT | Current proposals are ambiguous as they may have different understandings. We propose:   * The capability of the MNO to control the data transfer to the server for UE-side data collection. * The capability of the MNO to control the data transfer from the server for UE-side data collection. * The specific entity within the MNO to control the data transfer to and from the server for UE-side data collection. * The protocols ~~or~~ and methods utilized by the MNO to control the data transfer to and from the server for UE-side data collection. * The MNO's ability to manage (e.g., allow/disallow, initiate/terminate, prioritize/de-prioritize, etc.) the data transfer to and from the server for UE-side data collection. |
| Nokia | Yes, but it seems that the fourth bullet is a more descriptive version of the first bullet. The list could be consolidated by deleting the first bullet.  It should also be noted that there may be multiple protocols, e.g., for solution 3, there is an interface between OAM and the gNB, and between the gNB and the UE, each with a very different protocol. Furthermore, there would be an interface between the OAM and the server for UE-side data collection, which we need to determine to be proprietary or 3GPP-based. |
| Ericsson | Yes. At this stage we do not need to discuss which NW node is in charge of handling those 4 bullets. That is beyond RAN2 scope.  However, we believe that we should not only focus on controllability/management of the data transfer. Also whether the UE is allowed to start the data collection for the purpose of UE-side training is something that should be controlled and managed by the MNO/NW, i.e. the UE should start the data collection only if that is allowed by the MNO/NW.  So we propose adding the following bullet:   * The MNO's ability to manage/control the initiation/termination of the data collection for UE-side model training.   [Rapp1] is it covered by the bullet “The MNO's ability to manage (e.g., allow/disallow, initiate/terminate, prioritize/de-prioritize, etc.) the data transfer to and from the server for UE-side data collection”? |
| Huawei, HiSilicon | Yes with comments.  We have a high level question for clarification:  **At RAN2#125bis, it was agreed to defined what is control of data collection and waht is visibility of data content, and the necessity of controllability was not discussed before. Perhaps the email rapporteur could clarify whether/how we discuss the necessity of controllability in this email.**  For dimension discussions, we think the scope should be: capability of control and what to control.  1st bullet and 4th bullet should be about the capability of the MNO, so we suggest to merge them into a single bullet. In addition, regarding "ability to manage the data transfer", our view is that MNO should firstly be aware of the collected data, and then it can be able to control the data collection. So we think the discussion here is also relevant to section 2.5. **Our suggetion is: 1st/4th bullet is replaced by the following:**  **The capability of the MNO to control and manage the data transfer process, by considering the visibility of data content.**  for 2nd and 3rd bullet, we think they are not part of dimensions, and they are about details of controllability. **So we prefer to not list both bullets here.** If needed, we could address them later. |
| OPPO | Yes with comments  Basically, we agree the intention in general, but would like to further clarify the boundary between bullet 1 and bullet 4. My understanding is that bullet 4 is discussing how MNO can control the data collection task before the data is collected to the first termination entity while bullet 1 is trying to address the aspect how MNO can control the data sharing procedure after the data is collected to the first termination entity. If that is the case, the following revision suggestion can make it more clear:   * The capability of the MNO to control the data sharing procedure after the data is collected to the first termination entity, e.g. data sharing from the first termination entity to the server for UE-side data collection. * The specific entity within the MNO to control the data collection. * The protocols or methods utilized by the MNO to control the data collection. * The MNO's ability to manage (e.g., allow/disallow, initiate/terminate, prioritize/de-prioritize, etc.) the data collection task before the data is collected to the first termination entity. |
| Mediatek | I revised the bullets based on the comments received so far. We agree to discuss the above three bullets for controllability. |
| vivo | Beside above controllability aspects, It is essential to clarify the controllability for data collection at the server. We think the following aspects of controllability should be considered:   * Maintenance of user consent; * Selection of desired UEs to enable the server only collects data from specific UE(s); * Management of the session/connection between UE and termination entity;   Awareness of data content; |

Based on the feedback received so far, it seems necessary to clarify the level of controllability. As an initial step, we define the levels of controllability of MNO over data transfer to and from the server for UE-side data collection as follows:

* Full Control: 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. For example, the UE should start the data transfer only if that is allowed by the MNO/NW.
* Partial Control: The MNO has some degree of control over the data transfer but may be limited by certain factors such as agreements with third parties. For example, the UE can start the data transfer without involvement of MNO/NW as long as the tunnel is available.
* No Control: The MNO has no capability to influence or manage the data transfer.

For solution 1a), the MNO has no specific controllability for transfer of the collected data for UE-side data collection. For example, the collected data might be transferred together with other UE traffic without differentiation.

**Q4.2:** **Do companies agree that in solution 1a) MNO has no specific controllability for transfer of the collected data for UE-side data collection? Otherwise,** **please clarify any controllability that the MNOs have.**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | We agree its intention, but we think the term “specific controllability” is not clear:   1. We think NW implementation can already achieve NW controllability, e.g., by assigning a dedicated PDU session only for UE sided data collection. In this case, since the UE can’t mix other type of traffics with training dataset, it is at cost of radio resource inefficiency. 2. To alleviate the issue of radio resource inefficiency in 1), the controllability can be enhanced by limited spec change, e.g. SA2 introduce a new 5QI for dataset transfer of training. With it, the UE can mix other type of traffics with training dataset in one PDU session, and NW can differentiate it.   To make it clear, we suggest below rewording:  **in solution 1a) although it may not fully optimize system performance, MNO can already have ~~has no specific~~ controllability for transfer of the collected data for UE-side data collection via NW implementation.** |
| BT | Yes, but as mentioned before, 1a is a solution totally outside 3GPP. No need to discuss 1a |
| Nokia | Maybe. If the MNO knows the IP addresses of the OTT servers, traffic to the servers could be throttled. But there is, otherwise, no specific control over aspects of the data collection procedures. |
| Ericsson | Yes, if there is no SLA between MNO and the OTT server application.  However, if an SLA is present between MNO and OTT server, then the control is also possible and needed with 1a), i.e. the data stream destined to the IP address associated to the OTT server can be controlled/managed by the MNO as any other traffic.  So it should be either clarified that solution 1a) implies no SLA, or if with solution 1a) it is possible to have SLA, then the controllability is always possible. |
| Huawei, HiSilicon | Yes.  Solution 1a) cna work without specific controllability. |
| OPPO | Yes |
| Mediatek | Yes. No control from MNO over the data transfer from UE to the server. |
| vivo | Yes |

In solution 1b, it is recognized that the MNO retains a certain level of control over the data transfer to the server for UE-side data collection. The entity within the MNO responsible for this control might be a NF in the CN, such as the DCAF. The MNO exercises this control at the granularity of the PDU session, as governed by the SLA. However, since the initiation of the data collection procedure is triggered by the server as specific types of services, the MNO may not have the authority to directly control and manage the data transfer process.

**Q4.3: Related to solution 1b, can the MNO have control/management over the data collection for UE-side data collection, given that it may be managed by a NF within the CN, with control granularity at the level of the PDU session per SLA? Please clarify any other controllability that the MNOs might have in solution 1b.**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes, what RAN2 can clarify is the CN control, but the details of CN implementation can be discussed by SA2. |
| Apple | Yes with comments:  It is our understanding that from controllability perspective, the main difference of solution 1b from solution 1a is that it has one more control in CN via a NF (e.g. NWDAF, DCAF). However, the issue of radio resource inefficiency is same as solution 1, and the deployment of NF for data analysis (e.g. DCAF or NWDAF) in practical network is not successful so far.  Thus, to reflect above issue, we suggest below rewording:  **Related to solution 1b, although it may not fully optimize system performance, MNO can have control/management over the data collection for UE-side data collection, given that it may be managed by a NF within the CN, with control granularity at the level of the PDU session per SLA.** |
| BT | No. Considering Q3.2, our understanding is that UE-side data collected is inside the MNO. That means, the UE-side data collection belongs to the MNO. Therefore, it has control and awareness over collected data. |
| Nokia | Maybe. Depending on whether our interpretation of solution 1b matches what we have shown in our answer to Q3.2, the level of control could be considered to be more than just at the PDU session level.  However, if 1b is simply a server attached to a UPF and has an IP address internal to the network or specific to a DN for data collection, then we would agree with the PDU session level of granularity. |
| Ericsson | Yes.  The MNO can control and manage the entire data collection procedure via the UPF, and the gNB, as it would do for any other service injected into the 3GPP network. For example, the operator can configure a gNB/UPF with criteria to allow/disallow the transfer of data, it can control the amount of data transferred, it can control and manage the QoS of the data collection traffic, and the related overhead.  Based on SLA, the MNO can also control, e.g. by assigning different QoS flows, which data should terminated inside the MNO (i.e. the data that should be visible) and outside the MNO (i.e. the data that should not be visible). |
| Huawei, HiSilicon | No.  How MNO in involved in Solution 1b) is unclear to us. This is related to our comments in Q2.2 (inside MNO), and also related to the dimension discussion in Q4.1. |
| OPPO | Yes with comments  In general, we have the similar sense that a NF within the CN, e.g. DCAF, may be involved to control the data collection task at the level of the PDU session per SLA. But which node will generate/manage the data collection configuration is still unclear, in our view, at least the NF mentioned above can generate/manage the data collection configuration, which should be clarified further to better understand how solution 1b works in a whole picture. |
| Mediatek | Yes. We assume it’s partial control. The NW/MNO may not be able to start/stop the data transfer between UE and the server as long as the data tunnel is established. |
| vivo | Yes, but for solution 1b, the following controllability can be considered:   * Maintenance of user consent; * Selection of desired UEs to enable the server only collects data from specific UE(s); * Management of the session/connection between UE and the server; |

For solution 2, it is recognized that the MNO has full level of control over the data collection process. The entity within the MNO responsible for this control is a NF in the CN, such as NWDAF, LMF (for positioning), or other relevant NFs. There are two potential control methods:

* Option 1: The MNO may manage data collection through NAS signaling (or LPP for positioning).
* Option 2: Alternatively, the MNO may manage the data collection process from the CN to the UE via RAN nodes, utilizing RRC signaling.

This level of control enables the MNO to directly manage the data collection process as required.

**Q4.4:** **Do the companies agree that in solution 2, the MNO’s control over the data collection for UE-side data collection is characterized by full controllability? Please clarify whether it is managed by a NF within the CN through NAS signaling directly (option 1) or through RRC signaling via RAN node (option 2), or others?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes, what RAN2 can clarify is the CN control for this solution, but the details of CN implementation can be discussed by SA2.  In addition , we think the potential control methods for solution-2 can be also appliable to solution-1b.  From signaling perspective, we believe the NAS layer signalling should be prioritized. |
| Apple | Yes but we have below question for clarification for option 2:   * On RRC signaling, to help understanding, is it signaling like QoE (i.e. a transparent container in RRC message)?   [Rapp 1] It’s possible that there is certain interaction between CN and RAN, and it is RAN which controls the data collection process through RRC message. If it is a transparent container in RRC message, it is considered as a variation of option 1 via NAS signaling? |
| BT | Yes to "the MNOs have full control over data collection configuration and collected data for UE-side model training, management and inference on the CN”.  It is too early to decide the concrete signaling. |
| Nokia | Option 1: Maybe. It isn’t clear how the CN could provide detailed configurations to the gNB, though.  If the approach is CN-based, then the initial 3GPP-based control mechanism for DC for BM could come from the CN over NAS. NAS messages would be required to configure the UE for data collection for BM and NAS messages would be required to transport the measurements to the AMF and then further through the CN.  For positioning data collection, we agree that LPP would be the mechanism used to manage data collection from the UE.  Option 2: No. Currently, there is a mechanism to pass a signaling-based MDT Trace Activation from OAM to the gNB through the AMF, but we do not think that is the intention of this solution 2. We see this as a RAN-based or OAM approach. And for this option, the CN would still need to configure the RAN node for physical layer data collection. |
| Ericsson | Yes, but we would like however to clarify that the level of controllability is the same as in solution 1b, i.e. not clear why the term “full” controllability is used here. The only difference between the controllability of option 1b, and option 2, is that in option 1b the controllability is at UP level, whereas in option 2 it could be at CP level. Related to whether to use NAS or RRC signalling, we are concerned in general about CP solutions. NAS is the natural candidate for CN-centric approach, however that is not designed to carry large amount of data. Hence, the feasibility of using NAS for option 2 should assessed by SA2. |
| Huawei, HiSilicon | No.  The terminology "full controllability" is to be clarified. There has not been definitions for it.  We think it is too early to discuss the signaling details. In our understanding, we could discuss whether MNO has controllability over data collection process and data content at a high level.  For controllability over data collection process, if solution 2 is to use control plane for data collection, MNO should to able to have controllability of data collection process.  However, as we commented for Q4.1, for dimension discussions, we suggest to consider controlling data collection process and the visibility of data content together. For the data content aspect, we provide comments in section 2.7, and we do think discussion of controllability is inseparable from discussion of visiblity of data content. |
| OPPO | Comments  Actually, we don’t quite understand the meaning of ‘**full controllability’** mentioned in this question. My understanding is that Q4.2~Q4.5 should be answered on top of Q4.1, to differentiate the real controllability level among 4 solutions on the table, we should touch the specific aspects addressed in Q4.1; otherwise, any generic question does not make much sense for our discussion as we may still back to the beginning.  As for Option 1 and Option 2, we also don’t know what ‘manage’ refers to, does it refer to data collection configuration control or data collection reporting control or both? The answer can be totally different if we focus on different use cases, for BM use case, we think Option2, i.e. RRC visible signaling, should be the baseline no matter we’re talking about data collection configuration and data collection reporting; for positioning use case, for data collection configuration part, Option2, i.e. RRC visible signaling, may be possible while for data collection reporting part, Option 1, i.e. NAS signaling (LPP message) should be the baseline. |
| Mediatek | Yes, but  For solution 2, the MNO/NW can have full control of the data transfer to and from UE and the server for UE-side data collection if option 2 is considered. It requires coordination between CN and RAN to manage the data transfer procedure.  Same level of controllability, e.g., partial control as solution 1b is also possible for solution 2. |
| vivo | See comment:  We tend to focus on positioning for solution 2, i.e., CN collects data from UE/PRU via LPP procedures and further transfers of the data to the server.  In this case, the CN has full controllability over the data collection, including:   * Maintenance of user consent; * Selection of desired UEs to enable the server only collects data from specific UE(s); * Management of the session/connection between UE and termination entity; * Awareness of data content.   The LPP procedures can be categorized as Option 1. |

For solution 3, it is recognized that the MNO has full level of control over the data collection process. The entity within the MNO responsible for this control is OAM. The MNO controls the data collection process through RRC signaling via RAN node. This level of control enables the MNO to directly initiate and terminate the data collection process as required.

**Q4.5: Do the companies agree that in solution 3, the MNO’s control over the data collection for UE-side data collection is characterized by full controllability, managed by OAM through RRC signaling via RAN node?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes, what RAN2 can clarify is the OAM control for this solution, but the details of OAM implementation can be discussed by SA5. Meanwhile, it should be noted that the interface between OAM and RAN is non standardized interface.  It would be helpful to clarify how the OAM manage the data collection procedure through RRC signaling via RAN node.  We assume the OAM management can reuse the MDT framework. The OAM triggers the UE-side data collection via RAN (RRC signalling). The detail is different from MDT but the triggering and request the RAN node to do/signal something to the UE will be similar. We may consider this as baseline, then can discuss further, e.g. whether new framework is needed or not. |
| Apple | Yes |
| BT | Yes to "the MNOs have full control over data collection configuration and collected data for UE-side model training, management and inference on the OAM”. |
| Nokia | Yes. |
| Ericsson | Yes. We would like however to clarify that the level of controllability is the same as in solution 1b and 2, i.e. not clear why the term “full” controllability is used here. The only difference between the controllability of option 1b, and option 2/3, is that in option 1b the controllability is at UP level, whereas the in option 2/3 it is at CP level.  [Rapp1] I clarified the level of the full control. Please check. Similarly, we would also like to clarify the terminology “directly manage”. In option 1b, 2 and 3, there is a 3GPP network involvement, so in all these options there should be the possibility for the MNO to manage the data collection procedure. In option 1b, the management is done e.g. by an NF, in option 2 by a CN node, in option 3 by the OAM.  [Rapp1] This part is not important, so I removed it. In particular for the option 3, it should be clarified how the OAM can interact with the server for UE-side data collection (former OTT server) both for the case in which that is inside and outside the MN. That is important because the need for initiating/terminating a data collection session comes from the training entity. |
| Huawei, HiSilicon | No.  The terminology "full controllability" is to be clarified. There has not been definitions for it.  We think it is too early to discuss the signaling details. In our understanding, we could discuss whether MNO has controllability over data collection process and data content at a high level.  For controllability over data collection process, if solution 3 is to use control plane for data collection, MNO should to able to have controllability of data collection process.  However, as we commented for Q4.1, for dimension discussions, we suggest to consider controlling data collection process and the visibility of data content together. For the data content aspect, we provide comments in section 2.7, and we do think discussion of controllability is inseparable from discussion of visiblity of data content. |
| OPPO | Comments  Actually, we don’t quite understand the meaning of ‘**full controllability’** mentioned in this question. My understanding is that Q4.2~Q4.5 should be answered on top of Q4.1, to differentiate the real controllability level among 4 solutions on the table, we should touch the specific aspects addressed in Q4.1; otherwise, any generic question does not make much sense for our discussion as we may still back to the beginning.  More addition, we’d like to clarify that solution 3 is not suitable for positioning use case compared to solution 2. My understanding is that solution 3 and solution 2 can aim for different use cases. |
| vivo | See comments:  We tend to focus on beam management for solution 3 and reuse MDT framework for NW side data collection, i.e., OAM collects data from UE via MDT and further transfers the data to the server.  In this case, the OAM and gNB have full controllability over the data collection, including:   * Maintenance of user consent; * Selection of desired UEs to enable the server only collects data from specific UE(s); * Management of the session/connection between UE and OAM; * Awareness of data content. |

## 2.5 Visibility of data content in MNO

As a preliminary measure, the nature of the data content can be described by its format, type, value and others. Visibility refers to the extent to which the MNO is able to aware, access or even comprehend this data content.

Companies are invited to provide the definition on visibility and to what extent (aware, access or comprehend) the visibility is preferred.

**Q5.1 Companies are invited to clarify their understanding of data content visibility and specify their preferred level of visibility.**

|  |  |
| --- | --- |
| Company | Comment |
| NEC | In our understanding, the data content should be only visible by the termination point. |
| Apple | First, it is not clear whether the “data content visibility” could comply with regulations in different countries, e.g. GDPR. It is beyond 3GPP scope to define those requirements/mechanisms closely related based on legal grounds.  Second, assuming 3GPP is going to work on this requirement, we believe it is in remit of SA1 instead of RAN2. So, RAN2 should not define any requirement on behalf of SA1.  Third, regarding partial and fully data content visibility, SA3 should be consulted as there is potential impact on UE privacy and security and potential contradict with regional regulations.  Just to facilitate RAN2 discussion on clarifying solution details, we can compromise to start discussion on solution details based on **assumption of Rapporteur’s suggested 3 levels “visibility”**. However, we have strong concern that this assumption may be misused as requirement of “visibility”. Thus, **when capturing solution details, we requested to clearly clarify below points:**   1. **This is just an assumption for RAN2 study purpose, and the definition of “visibility” and the category of visibility should be discussed and defined in SA1.** 2. **RAN2 should not agree any definition and requirements on “visibility”.** 3. **Without proper definition and requirements from SA1 and guidance from SA3, RAN2 shall not conclude any solutions as it may contradict with regulations.** |
| BT | visibility for us mean awareness, access, and comprehend |
| Nokia | Visibility has at least two characteristics: knowledge that a certain type of data is being transmitted from the UE (e.g., by knowing that traffic to a certain IP address is always for training data collection); or knowledge about the contents of the data, e.g., the format of the data, or what a string of bits represents in terms of type and value.  We think that both characteristics are necessary to enable full visibility from the MNO. |
| Ericsson | OK with the definition given by the rapporteur. |
| Huawei, HiSilicon | In our understanding, visiblity of data content means each data to be collected should be standardized in relevant signalling level. It means that MNO should be able to see from the collected data what has been collected from UEs/users. |
| OPPO | My understanding of full data visibility has two levels of meanings:  Level 1: all collected data metrics are specified with open format, which means the data format is clearly defined in the spec.  Level 2: no vendor can get more info from the specified data.  For instance, RSRP metric collected via RRC, this data type is specified and the physically meaning is the same no matter which vendor gets this data.  My understanding of Partial data visibility has two levels of meanings:  Level 1: all collected data metrics are specified with open format, which means the data format is clearly defined in the spec.  Level 2: some vendor can get more info from the specified data.  For instance, some ID metric, e.g. *systemInformationAreaID*, broadcast via SIB1, this data type is specified but the physically meaning, e.g. which area this *systemInformationAreaID* serves for, is maintained by Operator or NW vendor itself, which is usually unknown by the UE vendor. From UE vendor point of view, *systemInformationAreaID* is not fully understood although this ID is logically used by the UE vendor to judge the validity of a SI message.  My understanding of no data visibility has one level of meanings:  Level 1: collected data metric is unspecified with proprietary format, which means the data format is not defined in the spec, something like a container. |
| vivo | Level 0: MNO entity is not aware of data collection procedure;  Level 1: MNO entity is aware of data collection procedure, and the collected data may be in string format as a container;  Level 2: MNO entity is aware of the type and value of collected data, that is, the data is with specified format and value.  From our perspective, Level 2 is preferred for solutions 1b/2/3. |

As an initial step, we define the levels of data content visibility within the MNO as follows:

* No visibility: MNO is not aware of the collected data and cannot access the data content.
* Partial visibility: The MNO is aware of the collected data, has limited access/comprehension to some elements of the data content, allowing for limited access.
* Full visibility: The MNO is aware of the collected data, has complete access to all aspects of the data content, enabling thorough comprehension.

For solution 1a), the MNO has no visibility of data content. For solution 1b), the MNO has no or partial visibility of data content depending on the SLA. For solution 2 and 3, the MNO is able to have full visibility of the data content.

**Q5.2: Do companies agree that in solution 1a) MNO has no visibility of data content for UE-side data collection?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | No.  In solution 1a), it is still possible that part of data content compliance with regulations can be visible to MNO via offline engineering between particular pair of MNO and UE vendor. Thus, we suggest below change:  **in solution 1a) MNO has no or partial visibility of data content compliant with regulations** **for UE-side data collection depending on offline engineering.** |
| BT | visibility for us mean awareness, access, and comprehend |
| Nokia | Yes. While the MNO might be able to infer something about traffic to a certain IP address, the contents of the data would be unknown to the MNO. |
| Ericsson | Yes |
| Huawei, HiSilicon | Yes with comments.  For full visibility, it means complete access to data content, without any conditions, so the wording "if needed" is unclear to us. In addition, with this wording "if needed", the boundary between partial visiblity and full visibility is unclear.  **So we suggest to remove "if needed".** |
| OPPO | Yes |
| Mediatek | Yes. |
| vivo | Yes |

**Q5.3: Do companies agree that in solution 1b) MNO has no or partial visibility of data content for UE-side data collection depending on SLA?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | No, we think the data content should be visible to the CN entity as solution 2.  [Rapp] This question is for solution 1b, not solution2. |
| Apple | Yes with comments:   1. We think the difference between solution 1a) and solution 1b) is just whether the offline agreement has 3GPP impact or not (i.e., offline engineering in solution 1a vs SLA in solution 1b). Please note that both solutions need offline negotiation between particular pair of MNO and UE vendor. And the efforts of offline negotiation should be equal for solution 1a) and solution 1b). 2. In the case of “MNO has partial visibility of data content”, which part of the data content can be visible to MNOs should be left to SA3 to decide, as it is impacting UE security and privacy.   And similar to Q5.3, we think the partial visibility of data content should also fulfill compliance with regulations. Thus, we suggest below change:  **in solution 1b) MNO has no or partial visibility of data content compliant with regulations for UE-side data collection depending on SLA** |
| BT | No. We consider UE-side data collector is inside the MNO network. Consequently, MNO has full visibility |
| Nokia | No. We think that no visibility, partial visibility, and full visibility are all options for solution 1b data collection. We think that partial visibility should also be defined.  We think that partial visibility means that some fields are decodable by the MNO and others are not decodable, e.g., proprietary measurement types. Additionally, we do not think that discussion about SLAs is in RAN2 scope, and we should instead discuss what type of access could be enabled from a third party. |
| Ericsson | No, visibility (full or partial) can be achieved for option 1b, as in the other solutions. The data that should be visible (e.g. based on an SLA agreement) can be first terminated within a termination entity inside the MNO (as per Q3.2), whereas data that should not be visible can be encrypted by the UE and not made visible. For example, upon receiving a service request from the UE for data collection, the operator can configure the network to setup different QoS flows, i.e. a QoS flow for the data that should be visible to the MNO, and a QoS flow for data that should not be visible. With the same approach, all data can be made visible, if needed. The UE will then encrypt data accordingly to ensure that certain data can be visible to the operator. |
| Huawei, HiSilicon | No.  In Solution 1b), how MNO can access the data content is unclear to us. |
| OPPO | No  We think for solution 1b, MNO has no visibility of data content for UE-side data collection from 3GPP perspective. If companies think some types of data are visible to MNO within 3GPP, please clarify what metric is visible to MNO and how MNO can be visible from spec perspective. We understand the SLA is for PDU session not for data content, please clarify. |
| Mediatek | Yes.  If the server for UE-side data collection is outside of MNO, MNO has no visibility of data content, since the data is transferred directly to the server of the MNO without stopping in the MNO network. If the server for UE-side data collection is inside of MNO, MNO is able to have partial visibility or full visibility of the data content according to the SLA. |
| vivo | For solution 1b, the MNO should be able to have full control of the data content/type exchanged between UE and server. |

**Q5.4: Do companies agree that in solution 2 and 3 MNO is able to have full visibility of data content for UE-side data collection?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | No, we think the data content should be visible to the CN entity for solution 2.  It is FFS for Solution-3, which may be up to the discussion at SA5. |
| Apple | No, it depends on whether/how the data content is 3GPP specified:   * If whole data content is UE vendor’s proprietary format, MNO has no visibility because MNO can’t comprehend it and thereby can’t use the data. * If some part of data content is UE vendor’s proprietary format (e.g. a container in RRC message) and other part is 3GPP specified format, MNO may have partial visibility.   + In this case, which part of the data content can be visible to MNOs should be left to SA3 to decide, as it is impacting UE security and privacy. * If whole data content is 3GPP format (e.g. specified IE in RRC message), MNO may have full visibility.   Since RAN1 is still discussing specification of data/metric for data collection, RAN2 can’t assume all data contents are 3GPP specified. If necessary, we can send LS to RAN1 for confirmation.    And similar to Q5.3 and 5.4, we think the partial/full visibility of data content should also fulfill compliance with regulations. Thus, we suggest below change:  **in solution 2 and 3 MNO ~~is able to have~~ has no, or partial, or full visibility of data content compliant with regulations for UE-side data collection depending on whether/how the data content is 3GPP specified.** |
| BT | No. We consider UE-side data collector is inside the MNO network. Consequently, MNO has full visibility  [Rapp1] So the answer is yes, i.e., the MNO has full visibility? |
| Nokia | Yes. However, a transparent container has been discussed in the past to enable a “partial visibility” solution, at least for solution 3. |
| Ericsson | No. Adopting solution 2 or 3 does not imply automatically that all data are visible by default. There might be also some data collected by the UE that are not standardized, and those data cannot be read, hence they are not practically visible. This should be always possible, irrespective of the solution adopted. |
| Huawei, HiSilicon | No.  It depends on what data MNO is to collect.  For Solution 2 and 3, if CP solution is considered and each data is standardized in relevant CP signalling, full visiblity can be achieved. For example, if MDT framework is used for solution 3, and L1 measurements can be collected via necessary enhancements in RAN2, then L1 measurements are full visible at OAM side and also at server side. |
| OPPO | Yes with comments  Based on our understanding in Q5.1, both full data visibility and Partial data visibility have full visibility of data content, i,e, specified data with open format, the difference is on whether all vendors can get the same info from the specified data. To protect UE proprietary info, UE may get more info from some specified data. We suggest to use the following revision to make it clear:  For solution 2 and solution 3, all collected data is specified with open format. |
| Mediatek | Yes. Standardization is not the sole method to achieve visibility of data content. Visibility can be granted to the MNO through business agreements, whereby the vendor discloses the data to the MNO that enters into such a business or cooperation contract. |
| vivo | Yes |

## 2.6 Protocol layer for Data Transfer

For solution 1a and 1b, the server for UE-side data collection receives data from the UE through the application layer, utilizing a UP tunnel for transmission. It should be noted that from user’s perspective, the data transfer from the UE to the server via UP tunnel is consider as traffic and is subject to data usage charges.

**Q6.1: Do companies agree that in solution 1a) and 1b) the data transfer from the UE to the server for UE-side data collection is through the application layer, utilizing a UP tunnel for transmission?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | Yes |
|  |  |
| BT | Yes.  This data will be considered as user plane. Therefore, it is the end customer who must pay for it (perhaps without being aware of it). |
| Nokia | Yes. Is it further implied, then, that the protocol for data transmission is proprietary, or are there standardized solutions, e.g., in the case of solution 1b? |
| Ericsson | Yes |
| Huawei, HiSilicon | Yes |
| OPPO | Yes |
| Mediatek | Yes |
| vivo | Yes |

For solution 2, the UE collects training data and transfers it to the CN, which then forwards the data to the UE-side server. As a basic approach, the data transfer from the UE to the CN can be conducted through the NAS layer using CP tunnel. The necessity for a UP tunnel to facilitate data transfer from the UE to the CN depends on the data volume, which is based on the requirements provided by RAN1. The feasibility of support a UP tunnel falls under the purview of SA2.

**Q6.2:** **Do companies agree that in solution 2, the baseline method for data transfer from the UE to the CN, is through the NAS layer, utilizing a CP tunnel for transmission?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes, the CP can be the baseline. On the other hand, RAN2 should keep the UP possibility on the table which is pending SA2 discussion. |
| Apple | Yes |
| BT | Not in a position to answer this question without knowing the data volume required for different AI/ML algorithms. |
| Nokia | Yes. To properly differentiate between solution 1b and 2, we think that it is necessary to restrict solution 2 to NAS signaling between the UE and AMF. It is FFS whether additional signaling would be required between the AMF and gNB. |
| Ericsson | No, agree with BT. Feasibility should be assessed in SA2. NAS is not designed today to carry large amount of data. Since data collection may imply the transfer of large amount of data, whether NAS is a feasible solution should be discussed in SA2. RAN2 cannot rule out non-NAS based solution. |
| Huawei, HiSilicon | Yes with comments.  We think Solution 2 can take CP for discussions first, but no need to discuss signalling details for now. |
| OPPO | Yes |
| Mediatek | Yes |
| vivo | We tend to focus on positioning for solution 2.  The data can be transferred from UE to LMF via LPP, which can utilize either CP or UP tunnel |

For solution 3, the UE collects training data and transfers it to the OAM, which then forwards the data to the UE-side server. As a basic approach, the data transfer from the UE to the OAM via RAN node can be conducted through the RRC layer using CP tunnel. The necessity for a UP tunnel to facilitate data transfer from the UE to the OAM depends on the data volume, which is based on the requirements provided by RAN1. The feasibility of support a UP tunnel falls under the purview of SA5.

**Q6.3:** **Do companies agree that in solution 3, the baseline method for data transfer from the UE to OAM via RAN node is through the RRC layer, utilizing a CP tunnel for transmission?**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| NEC | Yes |
| Apple | Yes |
| BT | Not in a position to answer this question without knowing the data volume required for different AI/ML algorithms. |
| Nokia | Yes. We should reuse the legacy approach. |
| Ericsson | Too early to say, agree with BT. For the reasons mentioned above related to NAS, the feasibility of other solutions should be assessed by SA2/SA5. |
| Huawei, HiSilicon | Yes with comments.  We think Solution 2 can take CP for discussions first, but no need to discuss signalling details for now. |
| OPPO | Yes |
| Mediatek | Yes |
| vivo | Yes, we tend to focus on beam management for solution 2 and reuse MDT. |

**Q6.4: For solution 2 and 3, interested company please share their view on solution details of UP tunnel solutions.**

|  |  |
| --- | --- |
| Company | Yes/No (Comment) |
| Apple | We are open to discuss UP option for solution 2 and 3. However, we are not sure how it works for now. According to current NR definition of UP tunnel, it is between UE and UPF, i.e. **the terminated entity is UPF in UL**. It seems to be conflicted with option 2 (3GPP terminated entity is one NF in CN) and option 3 (3GPP terminated entity is OAM). We believe a new Network Architecture and new protocol stack is not in Rel-19 scope. |
| Nokia | With regard to Apple’s comment that for the CN approach, the UPF is the termination point of UP traffic, we think that the use of the UPF should not be considered a CN approach in particular. The CN approach should terminate in a CN entity other than a UPF prior to any data transfer to a server which trains UE-side models. |
| Ericsson | Solutions other than NAS/RRC can be taken into account and assessed by SA WGs. |
| Huawei, HiSilicon | For UP solutions for solution 2/3, the collected data is totally invisible to the intermediate node(s) (e.g. CN, OAM), then how to protect privacy is a big problem. |
| Mediatek | For UP solution for solution 3, maybe we can have UP tunnel between UE and OAM. The data can be transferred to from UE to OAM through UP tunnel with RAN control. |
| vivo | SA2 already introduced the UP solution for LPP procedure between UE and LMF. |

## 2.7 Privacy Concerns

Privacy concerns in the communications ecosystem can vary significantly among different stakeholders, such as MNOs, network vendors, OEMs, and chipset vendors. Each entity has its own set of sensitive information that needs to be protected from unauthorized disclosure. Here are more detailed examples of privacy concerns for each stakeholder:

MNO Privacy Concerns:

* Network Information Disclosure: MNOs may inadvertently disclose sensitive network information, such as deployment strategies, network configurations, and performance metrics, to servers outside their network.
* Data Transfer Risks: When transferring data to external servers, there is a risk that the data may contain undisclosed information about the network or its users/customers. This could include subscriber identities, locations, website visited, phone calls, etc and other proprietary information that MNOs are obligated to protect.

Network Vendor Privacy Concerns:

* Sensitive Information Leakage: Network vendors may possess proprietary algorithms, system designs, and other intellectual property that are integral to their competitive edge. Unintentional disclosure of such information to third parties could undermine their market position and lead to potential legal issues.
* Implementation Details Exposure: The specific details of how network equipment is implemented, including software and hardware design, are crucial for maintaining the security and integrity of the network. If such information is disclosed, it could be exploited for malicious purposes or used by competitors to gain insights into the vendor's technology.

Chipset Vendor Privacy Concerns:

* Proprietary Technology Exposure: Chipset vendors develop specialized hardware and software that may contain trade secrets or patented technologies. here is a risk that shared information could be unintentionally disclosed to unauthorized parties, leading to privacy breaches.

OEM Privacy Concerns:

* User Information Disclosure: OEMs handle a vast amount of user data, including personal information, usage habits, and location data. There is a significant privacy concern if this information is disclosed to external entities without user consent, potentially violating privacy laws and damaging the OEM's reputation.

For all stakeholders, it is essential to implement robust data protection measures, such as encryption, access controls, and privacy policies, to mitigate these privacy concerns. Additionally, compliance with regulations like the General Data Protection Regulation (GDPR) and other local privacy laws is crucial for maintaining trust and legal standing in the market.

Companies are encouraged to clearly define and express their privacy concerns, even if these concerns may not be the primary focus of RAN2. There is a prevailing sentiment that companies have underlying fears that they have not been able to precisely define. This ambiguity regarding the specific nature of their concerns is not conducive to advancing our collective work on data collection. A well-defined awareness of the privacy issues at stake is important for facilitating meaningful progress in our study.

**Q7: What’s your privacy concerns on the data collection?**

|  |  |
| --- | --- |
| Company | Comment |
| Apple | As both UE vendor and chipset vendor, we have strong concerns on privacy from below aspects:   1. Concerns on UE Privacy:    * As UE vendor, we have strong concern on exposure of UE privacy to unknown 3rd entity. We believe it is always most sensitive issue to all UE vendors, and it may violate privacy laws and damage the OEM's reputation.    * The UE privacy at least includes UE Identities (SUPI, IP address, etc.), personal information, UE location, and UE trajectory information, etc.    * Thus, explicit UE consent for data collection is a must. Furthermore, we don’t think existing 3GPP specified UE consent mechanism/requirements in TS 33.501 Annex V (e.g. consent of MDT) is applied in this case because we understand it only specifies whether and how NFs in MNO domain check the user consent before processing the data. User consent mechanism for the scenario of sharing data outside of MNO domain has not been discussed in 3GPP yet. Without a proper mechanism on collecting user consent for sharing data outside of MNO domain, any solution will be risky to violate the user privacy as well as the regional regulations. 2. Concern on exposure of proprietary implementation to other vendors:    * It is common understanding that chipset vendor can have some proprietary implementation on top of 3GPP specification. And UE differentiation is one of most important reasons why 3GPP can build successful ecosystem so far. We believe any challenging of this principle will challenge success of 3GPP.    * As chipset vendor, we have strong concern on exposure of our proprietary implementation to other vendors, including other chipset vendors, UE vendors, NW vendors, MNOs, and any 3rd entity.    * Thus, we do not agree any solution which has potential risk of exposure of proprietary implementation to other vendors. And we believe not exposure of proprietary implementation to other vendors should be one of the most important requirements for UE data collection design. |
| BT | It is the operator who is under regulatory restriction. The lack of control and knowledge over collected data may result on an undesired exposure of our customers personal data. Most likely, it will conclude on penalties for operators as we will break regional regulations. In consequence, MNOs requires tools to stop it.  Option 1a is the only one that exempts operators from any responsibility, but that solution is totally outside 3GPP. It cannot be even considered as baseline.  Collected data will expose our network design and network strategy. Any disclaim of that information will have a negative impact. |
| Nokia | An additional privacy concern on Network Vendor that is not covered here is that radio topology and radio specific settings can be disclosed, and standard should not enable this. |
| Ericsson | All the concerns above can be taken into account. Agree with Nokia, about adding NW vendor privacy concerns.  It is important also to stress that irrespective of the solution used for transferring the collected data, the parties involved in the data collection transfer should always ensure the above privacy principles. This means that also for 1a), the above principles should be respected by the involved parties. |
| Huawei, HiSilicon | **For the above bullets, our comments are as below:**  For Network Vendor Privacy Concerns:  Suggest to add the following bullet:  **• Violation of user privacy regulation: The equipment of the network vendor may be used for collecting user’s data without getting approval/consent from the user in advance, and this behaviour may violate the local regulations and risks the sales of the equipment.**  For Chipset Vendor Privacy Concerns:  Add more description for Proprietary Technology Exposure:  **Another risk is that some sensitive data of a chipset vendor may be exposed to a second vendor without the knowledge of the chipset vendor.**  For OEM Privacy Concerns:  Suggest to add the following bullet:  **• Proprietary Technology Exposure: OEM vendors develop specialized hardware and software that may contain trade secrets or patented technologies. here is a risk that shared information could be unintentionally disclosed to unauthorized parties, leading to privacy breaches. Another risk is that some sensitive data of a OEM vendor may be exposed to a second vendor without the knowledge of the OEM vendor.**  **We have more comments for listed solutions:**  For Solution 1a), if user consent is used via implementation, we understand that all the user data are delivered within the user itself, with the 3GPP entities involved as transparent path, so we observe no privacy issues.  For Solution 1b), as we commented in section 2.4 and 2.5, this solution can not achieve "full" controllability and visibility, so there may be privacy issues.  For Solution 2 and 3, as we commented above, some sensitive data of a chipset/OEM vendor may be exposed to a second vendor without the knowledge of the chipset/OEM vendor, so there may be privacy issues. |
| OPPO | From UE vendor, i.e. OEM, point of view, all data generated from UE internally naturally belongs to the customer, no matter the data is coming from modem or camera or other software/hardware. UE vendor should have full visibility of any data ported from UE side no matter the data collection termination entity, e.g. UE cloud/other server not belonging to UE vendor/NW entity within 3GPP; otherwise, there is high risk to violate privacy laws and damage the OEM's reputation. |
| vivo | 1. User data exposure without user consent. And user consent should guarantee specified data collection from UE.  2. UE vendor-specific info exposed to other vendors. |

# 3 Appendix

**Table 1 Comparative analysis among different data collection approaches**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Aspects** | **1a) OTT (3GPP Transparent)** | **1b) OTT (Non-3GPP Transparent)** | **2. Transfer via Core Network** | **3. Transfer via OAM** |
| **Termination Entity** | UE-side OTT server | UE-side OTT server | CN | OAM |
| **Inside/outside MNO’s network** | Outside | Inside/Outside | Inside/outside | Inside/outside |
| **Transport 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 (FFS: UP tunnel) | CP tunnel (FFS: UP tunnel) |
| **Protocol layer for data transfer** | Application layer | Application layer | NAS or RRC layer (FFS: transport layer of UP tunnel) | RRC layer (FFS: transport layer of UP tunnel) |
| **Data Collection Controller** | UE-side OTT server | CN (certain NF, e.g., DCAF) | CN | OAM/RAN |
| **Control Granularity by NW** | NA, the OTT server can directly request data from the UE. | Coarse e.g., based on SLA | Finer (e.g., per NAS procedure) | Finer (e.g., per RRC procedure) |
| **RAN Intervene** | No | No | Possible | Yes |
| **Radio Resource Efficiency** | Low, as the network cannot optimize radio resource usage for data transfer | Low, as the network cannot optimize radio resource usage for data transfer | Higher, due to the possibility of RAN intervene | Maximum |
| **Data format** | Non-standardized | Non-standardized | Standardized/non-standardized | Standardized/non-standardized |
| **Network Awareness of the data Content** | No | No | Yes, if the data content is standardized or disclosed to MNO. | Yes, if the data content is standardized or disclosed to MNO. |
| **RAN configuration/condition acquisition** | Unclear whether the RAN configuration/condition acquired by the UE can be transferred to the UE-side OTT server. | Unclear whether the RAN configuration/condition acquired by the UE can be transferred to the UE-side OTT server and how the CN can control it with limited intervene. | Controlled by MNO | Controlled by MNO |
| **Spec Impact** | No | Limited | High | High |
| **Security and Privacy Risk** | High, managed by the OTT application, with potential risks if not 3GPP compliant | Lower, managed by the OTT application and NF based on SLA. | Minimum, NW can enforce security and privacy protection. | Minimum, NW can enforce security and privacy protection. |
| **Involved WGs** | No | SA2 | RAN2, SA2 | RAN2, SA2, SA5 |

# 4 Conclusion



# 5 Reference

1. [R2-2403967](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403967.zip) Discussion on Definition of Different UE Side Data Collection Solutions
2. [R2-2402364](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402364.zip) Data Collection for UE Side Model Training MediaTek Inc.
3. [R2-2403235](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403235.zip) Data Collection for UE-Side Model Training Interdigital Inc.
4. [R2-2403473](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403473.zip) Discussion on Data Collection for UE-side Model Training Futurewei Technologies
5. [R2-2403378](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403378.zip) Consideration on UE Side Data Colection ZTE Corporation,Sanechips
6. [R2-2403492](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403492.zip) AI/ML Data Collection Requirements T-Mobile USA, Verizon, Charter, NTT DOCOMO, Deutsche Telekom, Turkcell, BT, AT&T, Nokia, Telecom Italia, CMCC
7. [R2-2403230](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403230.zip) Discussion on UE Side Data Collection Qualcomm Incorporated
8. [R2-2402375](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402375.zip) Data collection for UE-side model training Samsung
9. [R2-2402962](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402962.zip) Data Collection for Training of UE-side AI/ML Models Nokia
10. [R2-2403567](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403567.zip) Discussion on UE side data collection China Unicom
11. [R2-2402171](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402171.zip) Data Collection for UE Side Model Training OPPO
12. [R2-2402302](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402302.zip) Consideration on UE side data collection CATT
13. [R2-2402316](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402316.zip) Data collection for UE side model training Xiaomi
14. [R2-2402342](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402342.zip) Discussion on UE side data collection Spreadtrum Communications
15. [R2-2402375](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402375.zip) Data collection for UE-side model training Samsung discussion
16. [R2-2402478](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402478.zip) Label Dataset Transfer for Positioning UE-sided model training Intel Corporation
17. [R2-2402489](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402489.zip) Discussion on UE side data collection vivo
18. [R2-2402669](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402669.zip) Data Collection for UE side Model training NEC
19. [R2-2402732](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402732.zip) Discussion on data collection for UE-sided model Lenovo
20. [R2-2402864](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2402864.zip) Discussion on UE-sided data collection Apple
21. [R2-2403022](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403022.zip) Discussion on data collection for UE-sided model training CMCC
22. [R2-2403122](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403122.zip) UE-side data collection Fraunhofer HHI, Fraunhofer IIS
23. [R2-2403163](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403163.zip) Discussion on UE-sided data collection for training Huawei, HiSilicon
24. [R2-2403230](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403230.zip) Discussion on UE Side Data Collection Qualcomm Incorporated
25. [R2-2403567](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403567.zip) Discussion on UE side data collection China Unicom discussion
26. [R2-2403573](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403573.zip) UE side data collection LG Electronics discussion
27. [R2-2403658](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403658.zip) Discussion on UE side data collection Indian Institute of Tech (M), IIT Kanpur
28. [R2-2403661](file:///C:\Users\panidx\OneDrive%20-%20InterDigital%20Communications,%20Inc\Documents\3GPP%20RAN\TSGR2_125bis\Docs\R2-2403661.zip) UE-side Data Collection Ericsson