3GPP TSG RAN WG1 #106-e R1- 210xxxx

e-Meeting, August 16th – 27th, 2021

Source: Moderator (Qualcomm)

Title: Summary on evaluation methodology for XR

Agenda Item: 8.14.2

Document for: Discussion and Decision

# Introduction

This contribution is a summary on the email discussion on evaluation methodology.

# Outcome of RAN1 #106-e

# Discussion

## Coverage evaluation

The following SLS-based evaluation methodology for XR coverage evaluation was discussed in RAN1#105-e.

|  |
| --- |
| **(Coverage evaluation methodology)** For XR/CG in DL or UL, coverage is defined to be the A-percentile point in CDF of Coupling gain for the “satisfied” UEs, with #UEs per cell = B, for a given XR application (AR/VR/CG) in a given deployment scenario (DU/InH/UMa)* A = [5], other value can also be reported
* FFS: Value of B, e.g. B = 1, capacity, etc.
* Note: Coupling gain for coverage evaluation is defined as the ratio of received and transmitted power measured in dB, and includes antenna gains, path loss, shadowing, indoor- or body loss, etc. Example of coupling gain can refer to TR 37.910.
 |

In RAN1#105-e, in addition to the above approach, the traditional LLS based methodology, e.g., what is used in the CE study/work item was addressed as an alternative by a few companies.

In contributions for RAN1#106-e, a clear majority of companies [3][4][5][7][8][9][10][11][12][13][14][15][16][17] propose to adopt an SLS evaluation methodology for XR coverage evaluation, while pointing out the following drawbacks of the LLS based approach: LLS would require non-trivial effort for (i) defining the new simulation cases and parameters, (ii) for updating simulator, and (iii) running simulation and obtaining results. In addition, a concern about the above SLS-based approach pointed out by a few companies is that it is not clear which value(s) of B would be used.

To address the above issue of the SLS-based approach, the following alternative SLS-based approach is proposed in [8].

|  |
| --- |
| * For each drop (or simulation),
	+ Randomly drop only one UE in the entire network (or in all the cells) that is associated with one of the 3 center cells (or gNBs), i.e., only one of the center gNBs is activated. This is equivalent to randomly locating one UE in a center gNB with the infinity ISD.
	+ Calculate coupling gain (detailed formula may follow TR 37.910 – to be discussed and confirmed in RAN1#106-e)
	+ Run SLS according to capacity evaluation methodology and determine whether the UE is satisfied or not.
* Definition of the XR coverage
	+ [X] % point of the CDF curve of coupling gain for the satisfied UEs for a given application.
* Coverage is evaluated only for Uma according to the above methodology.
 |

This methodology is to evaluate the maximum coupling loss/gain in a similar manner with the traditional link-level based approach. It is discussed in [8] that this approach would neither require much effort for simulator changes nor much effort for running simulations. The following results in [8] show that the maximum coupling gain of -140dB to -148dB for the simulated cases would be aligned with what is expected from the conventional LLS approach.



Pros and cons of the three options being discussed by companies are summarized in the next table.

|  |  |  |
| --- | --- | --- |
| **Option** | **Pros** | **Cons** |
| Option 1: Conventional LLS based approach (e.g., what is defined in Rel-17 CE study item)  | * Accustomed approach
 | * Require a significant RAN1 effort to define the detailed methodology
* Non-trivial effort to update LLS simulator
* Non-trivial effort to generate final result
 |
| Option 2:SLS based approach: B UEs per cell (what was discussed in RAN1@105-e) | * Insignificant additional effort to get the result on top of the required effort to get XR capacity result.
* Inter-cell interference impact to coverage can be natively captured in the result for multi-cell simulations.
 | * Unclear what number of UEs per cell would be appropriate.
* How to interpret the gap in the coverage result from the traditional LLS based approach
 |
| Option 3:SLS based approach: a single UE per network (e.g., what is proposed in [8]) | * Insignificant additional effort to get the result on top of the required effort to get XR capacity result.
* The nature of the result is the same as that of the traditional LLS approach.
 | * Inter-cell interference impact to coverage can be natively captured in the result for multi-cell simulations.
* Inter-cell interference impact to coverage cannot be dynamically evaluated (note: a constant interference margin may apply in the final result).
 |

1. **Please share your view on which of the following three methodologies for XR coverage evaluation should be chosen as the baseline, together with additional details as necessary.**

**Option 1**: Traditional LLS approach (e.g., what is used in Rel-17 CE study item)

If you prefer this approach, please provide detailed methodology as necessary.

**Option 2**: The SLS-based that was discussed in RAN1#105-e, as copied below.

* For XR/CG in DL or UL, coverage is defined to be the A-percentile point in CDF of Coupling gain for the “satisfied” UEs, with #UEs per cell = B, for a given XR application (AR/VR/CG) in a given deployment scenario (DU/InH/UMa)
* Coupling gain for coverage evaluation is defined as the ratio of received and transmitted power measured in dB, and includes antenna gains, path loss, shadowing, indoor- or body loss, etc. Example of coupling gain can refer to TR 37.910.

If you prefer this approach, please provide proposed values of A and B. Also, please feel free to provide other comments.

**Option 3**: The following SLS-based approach.

* For each drop (or simulation),
	+ Randomly drop only one UE in the entire network (or in all the cells) that is associated with one of the 3 center cells (or gNBs), i.e., only one of the center gNBs is activated. This is equivalent to randomly locating one UE in a center gNB with the infinity ISD.
	+ Coupling gain for coverage evaluation is defined as the ratio of received and transmitted power measured in dB, and includes antenna gains, path loss, shadowing, indoor- or body loss, etc. Example of coupling gain can refer to TR 37.910.
	+ Run SLS according to capacity evaluation methodology and determine whether the UE is satisfied or not.
* Definition of the XR coverage
	+ X %-tile point in the CDF curve of coupling gain for all the satisfied UEs
* Note: This methodology may be suitable only for UMa.

If you prefer this approach, please provide proposed value of X. Also, please feel free to provide other comments.

|  |  |
| --- | --- |
| **Company** | **Comment** |
| CATT | Option 2 – we had discussed this issue for a few meetings. We made preliminary agreements in RAN1#105-e. It does not make sense to start over again.  |
| vivo | Our first preference is Option 3 and second preference is option 2, since more discussion on detailed simulation assumptions for Option1 is required, especially considering there is only two meetings left. While for Option 2 and Option 3, all the agreed simulation assumptions for capacity evaluation can be reused. For Option 2, it has been fully discussed in RAN1#105 meeting. Hence, it would be less controversial if we can seek consensus on the values of A and B. In our opinions, A = 5, B = Capacity value and 1 can be adopted. For Option3, it is a simplified method of option 2 and trade-off between Option1 and Option2. It would be helpful if an example ISD value can be provided to reduce the workload to find out the appropriate ISD. In addition, the value of X can be the same as the value A in Option 2 i.e., X=5. |
| MTK | If RAN1 agrees to evaluate coverage on top of capacity/power, our first preference is Option 2 and second preference is Option 3. |
| ZTE, Sanechips | For option 1, it requires additional workload to align the simulation assumptions and efforts for the completely new simulations.For option 2, it requires least simulation efforts, but the parameter B is still controversial. For option 3, it also needs to run new simulations as the setting of activated center gNB is different from the capacity/power evaluations. Considering the simulation assumption/results of capacity and power among companies are quite diverse and the limited TU, if the simulation of coverage is needed, we prefer an option with least efforts. |
| Nokia, NSB | Agree with CATT that this issue has been discussed already for a while.As a way to move forward with this, we suggest trying to first agree that the SLS-based approach is used for coverage evaluation (currently covered by Option 2 and Option 3) and then continue discussing the details of this SLS approach.***Proposed agreement:****The SLS-based approach is chosen as a baseline for Coverage KPI.** *FFS: Discuss and agree on the details of the SLS-based approach for coverage study.*

While we understand that LLS-based study may provide some additional insights to the topic, we believe it would be: (i) very challenging to align the LLS coverage results to other KPIs in the TR, as well as, more importantly, (ii) we might not have sufficient time left in the coming 3GPP meetings to agree on the unified LLS methodology that demands extensive further discussions. According to our analysis of the submitted Tdocs, the majority of companies prefer to have the SLS-based approach applied for coverage evaluation. |
| Futurewei | Comments on Option 2It is very clear that the third option (Option 3) is newly proposed here as an alternative because of the concerns we and some other companies raised in the RAN1#105-e meeting in regard to the Option 2 which are 1) It is not clear what an appropriate number of UEs should be * Case with number of users =1 provide no additional information over geometry
* Case with number of users = capacity, includes the intercell and intracell interference aspects therefore there is no clear way to decouple the effects of interference and coverage. There is no calibrated capacity, so the results are very diverse and may not be reused for coverage.

2) Our sim results show that if you consider the CDF of satisfied and unsatisfied UEs and CDF of only satisfied UEs the results show no to little difference comparing to CDF of all UEs.Overall, this method is flawed and can provide little useful information about the coverage. Comments on Option 3Option 3 is proposed to solve the issues mentioned for Option 2, however the following are major concerns* This is the first time an approach such as Option 3 has been proposed. Discussion is need to clarify how it works and how it compares to baseline LLS. We propose this to be discussed carefully first to prove its feasibility. The CE SI would be the appropriate place to do so.
* Most companies have the LLS ready for CE study. All agreements made so far can be reused including Traffic model and PDB aspect. As an example, please refer to R1-2107087. It is also not clear why would we give time and effort for capacity and power evaluations but not for coverage.
* Option 3 requires changes in the SLS simulator including only centre base station is turned on, definition of ISD is non-existent, UE association, etc. This may require extensive efforts to change SLS.

To address concerns on Option 1Some companies concerned about the workload involved in using LLS for coverage evaluation. To our understanding, the concerns include the simulation time/effort and discussion on detailed assumptions. About the simulation time/effort, it should be comparable to other options in order to generate enough data points to derive statistics. Though some may argue that option 2 needs not to run simulation in addition to that of capacity evaluation, this is not necessarily true. For the case of B=1, additional simulation needs to perform for many independent drops. For the case of B=capacity, alignment between the companies on the capacity value (and detailed assumptions) is needed first.About the detailed assumptions, as shown in our contribution, most of the assumptions are just reusing the agreed assumptions for SLS evaluation. Only a small number of details, such as number of transmissions and TBS/MCS and MIMO layers for link simulation, may needs discussions considering the traffic model and PDB. We’ve provided details in our contribution how it may be done. Note that even these parameters can be left for company to decision just as in the case of SLS where these are up to the scheduler implemented there.Overall, most companies have the LLS ready after CE study SI and methodology is mature. The additional amount effort is well within the scope of this SI. Note that the main goal of this SI is to develop traffic model and evaluation methodology and we should not take short cut on deciding methodology. |
| QC | We prefer Option 3.Regarding options we have following comments: * Option 1 (LLS base approach): The methodology and definition of coverage is very well defined and studied. However, LLS approach has not much been discussed so far in XR SI, and therefore, in terms of workload, it is the highest among three options. To use LLS, we need to agree a new set of simulation assumptions which may require some sort of calibration as well. Furthermore, there could be companies may not even have available resources to do LLS based evaluation in addition to SLS. Given that we have very limited time and resources, option1 is not a viable option.
* Option 2 (multi Ues in multi-cell): The important consideration in XR coverage study is whether RAN1 can do any meaningful coverage study in the context of XR. We think the choice of B value is important in this aspect.
	+ If B=1, the max coupling loss observed is basically limited by ISD of the considered scenario, which means the max coupling loss of the satisfied UEs is not the real maximum coupling loss (MCL) that the XR service can be supported. In this sense, the observed MCL is not that meaningful.
	+ If B=capacity, due to larger number of UEs (limited resources per UE, increased interference, etc), the MCL of satisfied UE is smaller than the MCL of all Ues. The MCL of satisfied UE is no longer limited by ISD, but it is limited by configuration (# of Ues/cell, application data rate, fps, etc). Basically, choosing B=capacity requires RAN1 to define a new notion of XR coverage, which is necessarily different from traditional coverage definition. The resulting XR coverage is to be defined as MCL of satisfied Ues in a network with a given number of Ues / cell and application, etc.
* Option 3 (single UE per network): By construction, it is no longer ISD limited, and the resulting metric of max coupling loss of satisfied UE is very much in line with the traditional notion of coverage (defined and used in coverage study). Moreover, the additional workload is also minimal in the sense that we don’t actually need to agree new set of simulation assumption except (network layout). Most of the SLS assumptions already agreed in RAN1 can be re-used. Due to this reason, we think option 3 is most viable choice, allowing RAN1 XR coverage study for the remaining time period of R17. The outcome is also as meaningful / informative as traditional coverage study.
 |
| InterDigital | Our first preference is Option 2 followed by Option 3, given that Option 2 can potentially allow coverage evaluation to be done in parallel with capacity evaluation. We think Option 3 can be done optionally as a maximum limit to the achievable coverage performance. Both Option 2 and Option 3 will not add additional complexity given the ongoing discussions.  |

## Mobility evaluation

[3][4] [13][14] discuss evaluation methodology for mobility evaluation. Their proposals are summarized in the next table.

|  |  |
| --- | --- |
| vivo(2106630) | *Proposal 5: For XR/Cloud Gaming mobility evaluation, RAN1 focuses on evaluating capacity performance degradation due to handover procedure.****Proposal 6: For XR/Cloud Gaming mobility evaluation, the number of discarded or severely delayed packets due to interruption delay can be used as evaluation metric.*** |
| Samsung(2106918) | *System-level evaluation of mobility for XR devices is de-prioritized and XR mobility performance is captured analytically in TR 38.838.* |
| Ericsson(2107630) | [*Inter-cell mobility is evaluated analytically by describing the currently specified mobility procedures from an XR service point of view, relying on the agreed traffic models and user satisfaction criteria.*](#_Toc79150373) |
| Nokia(2107656) | **Proposal 3:** *RAN1 shall not to conduct advanced dynamic system-level simulations to assess the HOF and PP handover performance at this point of time.***Proposal 4:** *Conduct simple analytical study of the number of affected XR frames for the different agreed XR traffic models from HO interruption times, considering traditional HO, CHO, and DAPS (FR1 only). The XR TR 38.838 shall include a Table (e.g. ala the one Table 1) with the HO interruption times, as well as calculation of the number of effected XR frames from such interruptions. Based on that, simple conclusions can be drawn on how this will impact the XR QoS/QoE, including potential pointers for possible enhancements.***Proposal 5:** *The mobility KPI for the XR study is defined as the number of XR frames that have violated their PDB due to the HO interruption times, when considering traditional HO, CHO, and DAPS (FR1 only). The duration of the HO interruption time is to be calculated analytically by following the appropriate durations and processing times incorporated in the HO, CHO, and DAPS, as detailed in TS 38.133. An example of a possible table to be used for this study is given below as Table 1.**Table 1: Summary of HO interruption time for different HO methods as per the requirements in 3GPP TS 38.133 [3].*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Component** | **Description** | **Baseline HO** | **Conditional Handover (Late data forwarding)** | **Conditional Handover (Early data forwarding)** | **DAPS (FR1-FR1)** |
| 1 | RRC Reconfiguration procedure delay | 10 ms | 0 (UE can still receive from source cell while decoding) | 0 (UE can still receive from source cell while decoding) | **Downlink:**Up to Tinterrupt1 on source cell and Tinterrupt2 on target cell [3]For intra-frequency DAPS HO, Tinterrrupt1= 1 ms and Tinterrupt2 = 1ms (assuming same BWP)**Uplink:**CBRA: UL switch (sending new UL PDCP SDU and the non-acknowledged PDCP PDUs to target cell) occurs after MAC CE contention resolution is received. Target will forward the UL packets to UPF after ~ 10 ms (2 Xn messages for Handover Success+ SN Status Transfer). Total interruption can be up to 10 ms + Tinterrupt2 = 11 msCFRA = UL switch is performed when RAR is received. Additional delay for sending RRC Reconfiguration Complete which is 8 ms in FR1. Total interruption can be up to ~ 19 ms |
| 2 | Target cell search Tsearch | 0(if target cell is known) | 0(if target cell is known) | 0(if target cell is known) |
| 3 | UE processing time Tprocessing | 20 ms(upper limit for FR2) | 20 ms(upper limit for FR2) | 20 ms(upper limit for FR2) |
| 4 | Fine time tracking and acquiring full timing information of the target cell TΔ | 20, 10 ms on average(default value for SMTC period) | 20, 10 ms on average(default value for SMTC period) | 20, 10 ms on average(default value for SMTC period) |
| 5 | Tmargin (time for SSB post-processing) | 2 ms | 2 ms | 2 ms |
| 6 | Delay to acquire the first available PRACH in target gNB TIU | up to 20, 10 on average(for smallest value of x =1 defined in tables 6.3.3.2-2 and 6.3.3.2-3 of [4] for FR1 and table 6.3.3.2-4 for FR2) | up to 20, 10 on average(for smallest value of x =1 defined in tables 6.3.3.2-2 and 6.3.3.2-3 of [4] for FR1 and table 6.3.3.2-4 for FR2) | up to 20, 10 on average(for smallest value of x =1 defined in tables 6.3.3.2-2 and 6.3.3.2-3 of [4] for FR1 and table 6.3.3.2-4 for FR2) |
| 7 | PRACH preamble transmission | 1 slot(FR1/FR2: 1/0.125 ms) | 1 slot(FR1/FR2: 1/0.125 ms) | 1 slot(FR1/FR2: 1/0.125 ms) |
| 8 | UL Allocation + TA for UE | 10 slots(FR1/FR2: 10/1.25 ms) | 10 slots(FR1/FR2: 10/1.25 ms) | 10 slots(FR1/FR2: 10/1.25 ms) |
| 9 | UE sends RRC Reconfiguration Complete | 8 slots (FR1/FR2: 8/1 ms)  | 8 slots (FR1/FR2: 8/1 ms)  | 8 slots (FR1/FR2: 8/1 ms)  |
| 10 | Additional Interruption time from data forwarding | 0 ms | 10 ms (2 Xn messages for Handover Success+ SN Status Transfer) | 0 ms |
|  | **Ball-park total delay [ms] (FR1/FR2)** | **71/54.4 ms** | **71/54.4 ms** | **61/44.5 ms** |  |

 |

1. **Please share your comment on the following proposal on evaluation of XR mobility performance based on the contributions .**

**FL proposal on evaluation of XR mobility performance**

* XR mobility performance is evaluated analytically taking into account mobility procedures, agreed traffic models, and user satisfaction criteria.
	+ It is driven by contributions what/how to capture in the TR w.r.t. XR mobility evaluation. Companies are encouraged to submit mobility evaluation results, based on which RAN1 will discuss what/how to capture XR mobility performance in the TR.

|  |  |
| --- | --- |
| **Company** | **Comment** |
| CATT | Mobility aspects should be deprioritized in the evaluation. The mobility issue is the XR service degradation due to the long switch over time during the handover. The physical layer switch over and state synchronization during handover would take tens of milliseconds.  |
| vivo | We understand the intention of FL proposal. Given that there is limited time left for the SI and much work for capacity/power/coverage evaluation, evaluating mobility by analytical method and driven by companies would be the a more appropriate way. So we are supportive of FL’s proposal.We would like to clarify a bit more on how to evaluate mobility analytically. For mobility evaluation for XR, we can mainly focus on the impact of interruption due to handovers. And it would be too complicated to simulate the capacity performance by modelling the detailed handover procedure in SLS. Therefore, performance impacts due to handover procedures can be evaluated by numerical analysis, based on the assumptions of handover probability and interruption delay. The detailed evaluation steps are as follow:* **Step 1: Calculate handover probability**
	+ According to the typical topology scenario and UE speed, the handover probability can be calculated, e.g.

|  |  |
| --- | --- |
| **Parameters** | **Values** |
| Scenario | Reuse simulation assumptions as FeMIMO inter-cell mobility evaluation in R1-2007151.Dense Urban:Here X (in meter) is a uniformly distributed random variable U[26,34]. One UE is dropped and starts at P and moves along the 120-deg line downward to Q. |
| ISD(m) | 200 |
| X(m) | 26 | 34 | 26 | 34 |
| (a)Distance(P, Q)(m) | 492 | 501 | 492 | 501 |
| (b)UE speed(km/h) | 120 | 60 |
| (c)T(P, Q)(s),c=a/b | 14.76 | 15.03 | 29.51 | 30.05 |
| (d)Handover times | 7 |
| (e)Handover probability(times/s),e=d/c | 0.47 | 0.47 | 0.24 | 0.23 |

* **Step 2: Analyze interrupted packets for one-shot handover, e.g.**
	+ Assume handover interruption time is 40ms according to TS 38.133 and XR traffic periodicity is 60FPS, about 2.40 packets will be interrupted for one-shot handover.
* **Step 3: Calculate handover interrupted packets per second analyze the gap, e.g.**
	+ For 120km/h UE speed, the handover interrupted packets per second is about 0.47\*2.40 = 1.13 packets/s.
	+ Assume the PER requirement is 1% for the XR traffic (60FPS), the average packet loss per second cannot be larger than 0.6 packets/s, which means there is a huge gap for current handover mechanism to support XR traffic in high-speed case.
 |
| MTK | If RAN1 wants to evaluate mobility in Rel-17 XR, we agree with vivo that evaluating mobility by analytical method and driven by companies would be an appropriate way. The evaluation steps from vivo can be a starting point. |
| ZTE, Sanechips | Given the limited TU, we prefer to deprioritize mobility study in R17 XR SI. |
| Nokia, NSB | We agree, in general, with the main bullet. However, the proposed sub-bullet is confusing and requires further clarification. Is the intention to allow companies submit any mobility-related results without first agreeing on the unified methodology? If yes, then we don’t support this, as it may lead to a massive confusion at a later stage, when the results submitted by companies cannot be combined, as the assumptions and metrics are different.In order to move forward, we suggest the following modification to the proposed agreement:**FL proposal on evaluation of XR mobility performance*** XR mobility performance is evaluated analytically taking into account mobility procedures, agreed traffic models, and user satisfaction criteria.
	+ ~~It is driven by contributions what/how to capture in the TR w.r.t. XR mobility evaluation. Companies are encouraged to submit mobility evaluation results, based on which RAN1 will discuss what/how to capture XR mobility performance in the TR.~~
	+ FFS: Discuss and agree on a unified analytical evaluation methodology for Mobility KPI
		- NOTE: Strive to define the Mobility KPI during RAN1 106-e meeting.

We also see no technical reasons, why mobility evaluation should be deprioritized for the XR (especially, in comparison to coverage evaluation and/or extensive discussion on secondary-level parameters for **optional** traffic models).Mobility has been identified as one of the key KPIs in the SI description, along with capacity, power, and coverage. Moreover, some companies have already shown that coverage is not a critical factor in both baseline deployments (InH and DU). In contrast, companies have also indicated that there are notable mobility issues with time- and reliability-stringent XR services. Hence, we believe that mobility KPI deserves a full-scale discussion on the details of the analytical methodology to be used, striving to develop a simple analytical methodology already during the ongoing RAN1 106-e meeting.As an example, we suggest a very simple analytical definition for mobility KPI, where:Proposal: *The mobility KPI for the XR study is defined as the number of XR frames that have violated their PDB due to the HO interruption times, when considering traditional HO, CHO, and DAPS (FR1 only). The duration of the HO interruption time is to be calculated analytically by following the appropriate durations and processing times incorporated in the HO, CHO, and DAPS, as detailed in TS 38.133. An example of a possible table to be used for this study is given below as Table 1.*In case a more complex methodology (i.e., the one presented by VIVO) is preferred by the majority, the proposed definition can be well incorporated into Step 2, thus making the conclusions more reach and detailed. |
| Futurewei | Given limited time, prefer to deprioritize mobility in current SI |
| QC | We want leave mobility study up to each company. |
| InterDigital | We have similar understanding with MTK and Nokia on mobility evaluations. If RAN1 prioritizes mobility evaluation in SI, we support the proposal to evaluate XR mobility performance analytically.  |

## XR capacity for independent DL & UL simulations

[2][15] discuss how XR capacity is determined for the case when capacity for DL and UL is evaluated independently.

1. **Moderator’s view is summarized below. Please share your view and feel free to suggest an alternative.**

For the case when XR capacity is independently simulated, all results are separately captured in the TR. With that, it can be further discussed during the process of developing formal observations and conclusions which link (i.e., DL or UL) is the bottleneck for the given deployment scenario and configuration.

|  |  |
| --- | --- |
| **Company** | **Comment** |
| CATT | Since DL and UL XR evaluation are based on different XR applications respectively, the DL and UL evaluation results should represent the system capacity for each link. The bottleneck of system operation would be determined separated for either DL or UL too.  |
| vivo | We support the moderator’s view. |
| MTK | Support to develop formal observations and conclusions which link (i.e., DL or UL) is the bottleneck for the given deployment scenario and configuration |
| ZTE, Sanechips | Agree with moderator’s view. When DL and UL are evaluated independently, the channel quality varies in different link direction for the DL and UL simulation. For example, UE1 may have good channel quality in DL simulation but a worse channel quality in UL. Therefore, it's hard to obtain joint capacity on UE basis. Reporting DL capacity and UL capacity separately can help identify which link direction is the bottleneck. It seems joint capacity cannot provide additional information when DL and UL are evaluated independently.  |
| Nokia, NSB | Agree in general with the Moderator’s proposal. Results for DL and UL can be separately captured in the TR and later analyzed, when developing the formal conclusions (i.e., to identify which link is a capacity bottleneck or similar). |
| QC  | Agree with FL proposal. |
| InterDigital | We agree with the FL’s proposal |

## A clarification on power evaluation.

In [5], it is mentioned that the satisfied UE set used for statistical power saving gains needs to be clarified.

* Option 1: Collecting the satisfied UE set of the baseline and different UE power saving schemes PSS1, PSS2 etc., which are represented as S\_bl, S\_pss1, and S\_pss2 etc., Then the power consumption of the S\_bl in the baseline, the power consumption of S\_pss1 under PSS1, the power consumption of S\_pss2 under PSS2 etc. are collected. Finally, power saving gain of different power saving schemes is calculated using power consumption of different satisfied UE set.
* Option 2: Collecting the satisfied UE set of the baseline which is represented as S\_bl. Then the power consumption of the S\_bl under baseline and different power saving schemes is collected. Finally, the power saving gain of different power saving schemes is calculated using power consumption of same satisfied UE set.

[5] proposes that the satisfied UE set considered for UE power saving gain is satisfied UE of baseline, i.e. option 2.

1. **Per moderator’s understanding Option 1 is a common understanding in RAN1. The set of satisfied UEs can be different for different power saving schemes with different configurations (e.g., CDRX parameters). In the case when the power saving gain of a power saving technique is calculated only for the satisfied UEs, the set of satisfied UEs should be the one for the evaluated power saving technique. Please share your understanding/view.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
| CATT | The UE power gains of satisfied UEs need to further clarified. All UEs should be configured and operated with same power saving scheme in order to evaluate the power saving gain for the power saving scheme. The power consumption of all UEs should be computed during the simulation. If we only collect the power consumption of UEs satisfied the XR service requirements, we might miss the variation of UE power consumptions, e.g., large UE power consumption of PDCCH monitoring for XR packet late arrival by non-satisfied UE. We prefer to calculate power saving gain from all UEs instead of satisfied UEs. |
| vivo | We support the moderator’s view. |
| MTK | This topic seems to be discussed before. We prefer to calculate power saving gain from all UEs instead of satisfied UEs, while at the same time we do not see a large difference in the results derived from either all UEs or satisfied UEs. We are open to align the assumption if majority of companies think this issue is important. |
| ZTE, Sanechips | We agree that the satisfied UE should be considered for the evaluation of power saving gain. However, the cons of option1 and option 2 are:Option 1 may use different UE sets to calculate the power saving gain. Option 2 may include unsatisfied UEs to determine UE power saving gain. We think the satisfied UE sets should remain the same when calculate the power saving gain for a power saving scheme. Therefore, we think the satisfied UE set considered in UE power saving gain calculation is the intersection of the satisfied UE set in the baseline and the counterpart in a power saving scheme. A suggested Option 3 is:Option 3: Collecting the satisfied UE set of the baseline and different UE power saving schemes PSS1, PSS2 etc., which are represented as S\_bl, S\_pss1, and S\_pss2 etc., Then determining the intersection of the satisfied UE set in the baseline and different UE power saving schemes PSS1, PSS2 etc, which are represented as I\_b1\_pss1(=S\_bl∩S\_pss1), I\_b1\_pss2(=S\_bl∩S\_pss2).Then the power consumption of the I\_b1\_pss1, I\_b1\_pss2 in the baseline, the power consumption of I\_b1\_pss1 under PSS1, the power consumption of I\_b1\_pss2 under PSS2 etc. are collected. Finally, power saving gain of different power saving schemes is calculated using power consumption of different satisfied UE intersections. |
| Nokia, NSB | We agree with the Moderator’s understanding. The list of satisfied UEs may differ depending on (i) if any PSS is used; and (ii) which particular PSS is used. Hence, Option 2 mat be controversial, as the UEs in S\_bl may be both “satisfied” for “Always ON” and “non-satisfied” for some PSS schemes.Responding to CATT and MTK concerns, we agree that collecting only the PSS for satisfied UE may provide limited information. As we recall, that was one of the reasons, why it has been already agreed during RAN1 104-bis-e that:Agreement: For XR power evaluation (including baseline and power saving schemes), **companies report** **both** Option 1 and Option 2 results for evaluating the power saving gain.* Option 1: all UEs are considered
* Option 2: satisfied UEs only are considered

Hence, the statistics for all the UEs is also to be collected that should facilitate accounting for non-satisfied UEs. Moreover, comparing the PSS gains for “all UEs” with the PSS gains for “satisfied only UEs” may lead to some important observations (i.e., those that you listed, related to the increase in the UE power due to late packets arrival, etc.). Hope that this approach is sufficiently broad. Please, feel free to comment further if we didn’t fully understand your concern.Following the prior agreement listed above, **we suggest removing the word “only” from the second sentence** (in case that this text goes to any captured agreements/Moderator’s understanding/etc.), just to avoid possible confusion:**In the case when the power saving gain of a power saving technique is calculated ~~only~~ for the satisfied UEs, …**  |
| QC | We agree with the Moderator’s understanding. |

## Others

1. **Please feel free to discuss topics that are not discussed above.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
|  |  |
|  |  |

# Summary of contributions for RAN1#106-e

## Coverage evaluation

Companies’ views on coverage evaluation for XR are summarized as below.

|  |  |
| --- | --- |
| Huawei(2106457) | ***Observation 1: For coverage evaluation methodology, the LLS based and SLS based approach have drawbacks including large amount of additional workload, it’s unclear whether the methodology is accurate or meaningful for XR, whether there are any XR-specific issues, etc. Companies can further study coverage evaluation methodology if needed.*** |
| vivo(2106630) | *Proposal 2: For SLS approach for coverage evaluation, edge user coupling gain with #UEs/cell equal to capacity can be used as evaluation metric.**Proposal 3: For LLS approach for coverage evaluation, MIL (max isotropic loss) can be used as evaluation metric.**Proposal 4: For XR/Cloud Gaming coverage evaluation, both SLS and LLS approach can be considered, and it is up to companies to choose one or both of them.* |
| Samsung(2106918) | P**roposal 1***XR coverage system-level evaluations use A=5, re-use PDB assumptions from capacity evaluations and assume low-load, e.g. B = 1.* |
| CATT(2106950) | ***Observation 1: Adopting a definite low system load seems more reasonable for XR coverage evaluation.******Proposal 1:*** ***Low system load, e.g., 1UEs per cell, is mandatory for coverage methodology evaluation.*** ***Proposal 2: For XR coverage evaluation, more metrics should be reported together with A-percentile point in CDF of Coupling gain for the “satisfied” UEs, e.g., UE satisfied rate.***  |
| FUTUREWEI(2107087) | ***Observation 1:*** Under the simulation assumptions made in Section 2.1, an SNR of -3.37dB is required for target performance of 0.001 PER for 1-layer PDSCH transmission. For the case of two downlink layers, the performance is enhanced by 1.5dB.***Observation 2:*** The MCL for single layer transmission PDSCH is -142.37dB while the MCL for the case of two layers PDSCH transmission is -143.8dB.***Observation 3:*** Under the simulation assumptions made in Section 2.1, an SNR of 1.10dB is required for target performance of 0.001 PER for 2-layer PUSCH transmission for uplink traffic of 10 Mbps and an SNR of 0.0 dB is required for uplink traffic of 0.2 Mbps. ***Observation 4:*** For the case of 10 Mbps data rates, the MCL for single layer transmission PUSCH is 114.44dB. The MCL for two layers of transmission is 114.59. For the case of 0.2 Mbps data, rates the MCL 127.45 assuming CP-OFDM***Observation 5:*** With the assumption of B=capacity UEs/cell, the 5%-tile coupling loss obtained from considering satisfied UEs only compared to the 5%-tile coupling loss obtained from considering all UEs* 2 dB difference for DU scenario
* almost equal for UMa

Regarding obtaining coupling loss CDF for uplink channel, it is not clear what the value of B should be which is similar problem to that discussed in the downlink. Moreover, in case the value of B to be considered is equal to the capacity, it is further not clear whether the UL and DL joint capacity should be considered or not. ***Observation 6:*** For uplink channel coupling loss/gain results, it is not clear what the value of B should be. In the case of B=capacity, it is further not clear whether capacity is obtained from joint UL and DL simulations or based on uplink simulations only.***Observation 7:*** The assumption of having 1 UE per cell may not provide meaningful results as it is equivalent to a simple CDF of coupling loss.***Observation 8:*** With the assumption of having number of UEs per cell > 1 and specifically equal to capacity, interference affects the system capacity and thus it is not clear how one may analyze the effects of coverage separately***Observation 9:*** Capacity results have not been aligned between companies thus the use of B=capacity may be problematic.***Observation 10:*** The system level approach does not provide information on which channel or data stream is the bottleneck for coverage. ***Observation 11:*** Progress should not be pursued at cost of having accurate results. Moreover, resolving the issues mentioned above may cause a delay in SI progress as a lot of studies needs to be done for proving accuracy.***Observation 12:*** Link-level based coverage analysis is mature and more accurate and does not have the issues listed above as has been followed by CE and RedCap SIs in Rel-17.***Observation 13:*** The group may discuss the suitable parameters for link-level based coverage evaluations:* Size of packet: maximum size frame size according to the adopted statistical model
* Number of Retransmission with the agreed TDD configuration and PDB
* Receiver interference density to model the interference aspect

Based on the above analysis and observations, we have the following proposal:***Proposal 1:*** The link-level based evaluation methodology is adopted as evaluation methodology for XR coverage analysis.  |
| OPPO(2107280) | ***Proposal 1: If system evaluation method is applied for coverage evaluation, the following details are suggested:**** ***The smallest coupling gain for the “satisfied” UEs, i.e. A=0 can be used to identify the coverage gap between UL and DL;***
* ***UE number per cell should equal to capacity;***
* ***Other detail parameters for system evaluation for coverage can refer to parameters for system evaluation for capacity directly.***
 |
| Qualcomm(2107375) | ***Proposal 1: Use system level simulation (SLS) that has been used for capacity simulation for Rel-17 XR coverage study.******Observation 1:**** ***Option 1 (one UE in the entire network) covers the transition range of the coupling gain where the UE starts to become unsatisfied***
* ***Due to the limited cell size even in the UMa scenario, Option 2 (one UE per cell) does not cover the extreme condition for the UE to become unsatisfied as the coupling gain decreases and hence does not reflect the difference between all UEs and satisfied UEs, CG and VR, DL and UL. This will become highlighted for lower data rate such as CG8Mbps.***
* ***UL coverage is worse than DL***

***Proposal 2: Adopt Option 1 of the following options for XR coverage evaluation*** * ***Option 1: Only the center base stations are enabled; a single UE is randomly dropped in the entire topology for each run of the simulation***
* ***Option 2: All base stations are enabled; a single UE is randomly dropped in each cell for each run of the simulation***

***Proposal 3: Coverage is defined as the 1%, 5% or 10% points of the CDF curve of the coupling gain for the satisfied UEs*** |
| LG(2107462) | ***Proposal 1: Further discuss the objective of coverage evaluation for XR and the coverage evaluation methodology with the following note from the Chairman’s Notes in RAN1#105-e as a starting point:*****For companies to further study and if necessary, discuss in RAN1#106-e****(Coverage evaluation methodology)** For XR/CG in DL or UL, coverage is defined to be the A-percentile point in CDF of Coupling gain for the “satisfied” UEs, with #UEs per cell = B, for a given XR application (AR/VR/CG) in a given deployment scenario (DU/InH/UMa)* + A = [5], other value can also be reported
	+ FFS: Value of B, e.g. B = 1, capacity, etc.
	+ Note: Coupling gain for coverage evaluation is defined as the ratio of received and transmitted power measured in dB, and includes antenna gains, path loss, shadowing, indoor- or body loss, etc. Example of coupling gain can refer to TR 37.910.

An alternate method could be to use the “traditional” method such as what is used in the CE study/work item.For the #UEs per cell (B) for coverage evaluation, we can consider both B = 1 and B = capacity as a starting point for further discussion in RAN1#106-e meeting.  |
| MediaTek(2107501) | ***Observation 1: The “traditional” method such as what is used in the CE study/work item uses the link level simulation (LLS) which assumes a one-to-one communication. In this case, the coverage evaluation results would be quite hard to be utilized with the SLS results generated for capacity and power evaluation.******Proposal 2: Do not use the “traditional” method such as what is used in the CE study/work item since the LLS results would be quite hard to be utilized with the SLS results generated for capacity and power evaluation.******Observation 2: The coupling gain in 37.910 is only defined for DL, not UL. In 37.910, coupling gain is always evaluated with DL geometry.******Proposal 3: The “A-percentile point in CDF of Coupling gain” coverage evaluation is only applied to DL. The coverage evaluation for UL needs to be further studied.******Observation 3: If setting “B = 1”, most likely all UEs are satisfied, and the A-percentile coupling gain would be limited by the considered deployment scenario. This would not be that meaningful or informative.******Proposal 4: Setting “B = capacity” to provide SLS based coverage study of ”Coverage in the capacity regime”.*** |
| InterDigital(2107535) | **Proposal 1:** For XR coverage evaluation, reuse the methodologies and simulation assumptions agreed for system level simulation to evaluate capacity.**Proposal 2:** For XR coverage evaluation, apply the same per-UE satisfaction criteria (e.g. X%, PDB) used for capacity evaluation.**Proposal 3:** Adopt the definition for DL/UL coverage evaluation in XR/CG as stated in [1].**Proposal 4:** The value for A (90, 95, 99) percentile, can be reported by companies and as a baseline, B can be set to 1 while optionally, B can be chosen to be equal to the number of satisfied UEs. |
| Intel(2107617) | ***Proposal: For XR/CG in DL or UL, for coverage evaluation, companies can report the CDF of Coupling gain for the “satisfied” UEs, with #UEs per cell = 1, for a given XR application (AR/VR/CG) in a given deployment scenario (DU/InH/UMa)**** ***Note: Coupling gain for coverage evaluation is defined as the ratio of received and transmitted power measured in dB, and includes antenna gains, path loss, shadowing, indoor- or body loss, etc. Example of coupling gain can refer to TR 37.910.***
 |
| Ericsson(2107630) | Proposal 1 Reuse the end-user satisfaction criteria agreed for the capacity evaluations also for the coverage evaluations.Proposal 2 Coverage is defined as the probability that a user is satisfied when the number of users in the system is very low.Proposal 3 Use B=1 and A=5 to evaluate coverage for the various XR services. |
| Nokia(2107656) | **Observation 1:** *Option 1 (B=arbitrary value, B>1) is non-preferrable due to its insufficient flexibility.***Observation 2:** *Option 2 (B=1) and Option 3 (B=capacity) present two different approaches of defining the coverage KPI, where Option 2 defines it for the coverage-limited regime, while Option 3 defines it for the system that is limited by both coverage and interference constrains.***Observation 3**: *Option 2 (B=capacity) is non-preferrable due to its focus on interference-limited regime, complexity, and anticipated difficulty in converging the results among the companies.***Observation 4:** *Adopting B=1 creates the least inter-cell interference and coverage is typically defined for such cases where inter-cell interference is not accounted for.*The following **proposals** have been made:**Proposal 1:** *Adopt the proposal above as a starting point for defining the coverage KPI and continue discussing whether any further tuning and/or changes are needed.***Proposal 2:** *Adopt B=1 as a baseline for coverage KPI, as it makes the metric clearly coverage-limited (which is the exact meaning of a coverage-related KPI), easy to calculate (short simulations with B=1) and to align among companies (B=1 does not give a room for deviations in the modeled setup).* |
| Apple(2107769) | **Proposal 2:** Focus the XR study on system capacity and UE power consumption.For coverage study,system level evaluation is used. |
| DOCOMO(2107887) | **Proposal 4:*** *Confirm the proposal for coverage evaluation where B can be equal to system capacity.*
 |
| Xiaomi(2107906) | **Proposal 1: For coverage evaluation of XR services, the following options can be considered:****- Option 1: system level evaluation methodology with value of B = capacity****- Option 2: link level evaluation methodology in Rel-17 CE SI/WI** |

## Mobility Evaluation

Companies’ views on mobility evaluation for XR are summarized as below.

|  |  |
| --- | --- |
| Huawei(2106457) | For mobility evaluation, it seems there are no XR-specific issues. Therefore, we suggest to postpone or even not study mobility in R17 XR SI. |
| vivo(2106630) | *Proposal 5: For XR/Cloud Gaming mobility evaluation, RAN1 focuses on evaluating capacity performance degradation due to handover procedure.**Proposal 6: For XR/Cloud Gaming mobility evaluation, the number of discarded or severely delayed packets due to interruption delay can be used as evaluation metric.* |
| Samsung(2106918) | **Proposal 2***System-level evaluation of mobility for XR devices is de-prioritized and XR mobility performance is captured analytically in TR 38.838.* |
| OPPO(2107280) | Proposal 2: The evaluation on the impact of motility events on XR/CG is de-prioritized.  |
| Ericsson(2107630) | Observation 1 Compared to eMBB, XR applications will be more impaired by intra-cell and inter-cell mobility.Proposal 2 Inter-cell mobility is evaluated analytically by describing the currently specified mobility procedures from an XR service point of view, relying on the agreed traffic models and user satisfaction criteria. |
| Nokia(2107656) | **Proposal 3:** *RAN1 shall not to conduct advanced dynamic system-level simulations to assess the HOF and PP handover performance at this point of time.***Proposal 4:** *Conduct simple analytical study of the number of affected XR frames for the different agreed XR traffic models from HO interruption times, considering traditional HO, CHO, and DAPS (FR1 only). The XR TR 38.838 shall include a Table (e.g. ala the one Table 1) with the HO interruption times, as well as calculation of the number of effected XR frames from such interruptions. Based on that, simple conclusions can be drawn on how this will impact the XR QoS/QoE, including potential pointers for possible enhancements.***Proposal 5:** *The mobility KPI for the XR study is defined as the number of XR frames that have violated their PDB due to the HO interruption times, when considering traditional HO, CHO, and DAPS (FR1 only). The duration of the HO interruption time is to be calculated analytically by following the appropriate durations and processing times incorporated in the HO, CHO, and DAPS, as detailed in TS 38.133. An example of a possible table to be used for this study is given below as Table 1.**Table 1: Summary of HO interruption time for different HO methods as per the requirements in 3GPP TS 38.133 [3].*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Component** | **Description** | **Baseline HO** | **Conditional Handover (Late data forwarding)** | **Conditional Handover (Early data forwarding)** | **DAPS (FR1-FR1)** |
| 1 | RRC Reconfiguration procedure delay | 10 ms | 0 (UE can still receive from source cell while decoding) | 0 (UE can still receive from source cell while decoding) | **Downlink:**Up to Tinterrupt1 on source cell and Tinterrupt2 on target cell [3]For intra-frequency DAPS HO, Tinterrrupt1= 1 ms and Tinterrupt2 = 1ms (assuming same BWP)**Uplink:**CBRA: UL switch (sending new UL PDCP SDU and the non-acknowledged PDCP PDUs to target cell) occurs after MAC CE contention resolution is received. Target will forward the UL packets to UPF after ~ 10 ms (2 Xn messages for Handover Success+ SN Status Transfer). Total interruption can be up to 10 ms + Tinterrupt2 = 11 msCFRA = UL switch is performed when RAR is received. Additional delay for sending RRC Reconfiguration Complete which is 8 ms in FR1. Total interruption can be up to ~ 19 ms |
| 2 | Target cell search Tsearch | 0(if target cell is known) | 0(if target cell is known) | 0(if target cell is known) |
| 3 | UE processing time Tprocessing | 20 ms(upper limit for FR2) | 20 ms(upper limit for FR2) | 20 ms(upper limit for FR2) |
| 4 | Fine time tracking and acquiring full timing information of the target cell TΔ | 20, 10 ms on average(default value for SMTC period) | 20, 10 ms on average(default value for SMTC period) | 20, 10 ms on average(default value for SMTC period) |
| 5 | Tmargin (time for SSB post-processing) | 2 ms | 2 ms | 2 ms |
| 6 | Delay to acquire the first available PRACH in target gNB TIU | up to 20, 10 on average(for smallest value of x =1 defined in tables 6.3.3.2-2 and 6.3.3.2-3 of [4] for FR1 and table 6.3.3.2-4 for FR2) | up to 20, 10 on average(for smallest value of x =1 defined in tables 6.3.3.2-2 and 6.3.3.2-3 of [4] for FR1 and table 6.3.3.2-4 for FR2) | up to 20, 10 on average(for smallest value of x =1 defined in tables 6.3.3.2-2 and 6.3.3.2-3 of [4] for FR1 and table 6.3.3.2-4 for FR2) |
| 7 | PRACH preamble transmission | 1 slot(FR1/FR2: 1/0.125 ms) | 1 slot(FR1/FR2: 1/0.125 ms) | 1 slot(FR1/FR2: 1/0.125 ms) |
| 8 | UL Allocation + TA for UE | 10 slots(FR1/FR2: 10/1.25 ms) | 10 slots(FR1/FR2: 10/1.25 ms) | 10 slots(FR1/FR2: 10/1.25 ms) |
| 9 | UE sends RRC Reconfiguration Complete | 8 slots (FR1/FR2: 8/1 ms)  | 8 slots (FR1/FR2: 8/1 ms)  | 8 slots (FR1/FR2: 8/1 ms)  |
| 10 | Additional Interruption time from data forwarding | 0 ms | 10 ms (2 Xn messages for Handover Success+ SN Status Transfer) | 0 ms |
|  | **Ball-park total delay [ms] (FR1/FR2)** | **71/54.4 ms** | **71/54.4 ms** | **61/44.5 ms** |  |

 |
| DOCOMO(2107887) | **Proposal 1:***It is suggested that mobility evaluation is conducted in this study item to see the performance and whether any enhancement on mobility is needed for XR services.***Proposal 2:*** *The following mobility speed can be considered for XR mobility evaluations:*
	+ *Pedestrian (e.g. 3 km/h), vehicular (e.g. 60 km/h), and HST (e.g. 300 km/h or 500 km/h)*

**Proposal 3:*** *Further discuss whether analytical evaluation or system level evaluation based on Rel-17 MIMO mobility study is used for XR mobility evaluation.*
 |
| Xiaomi(2107906) | **Proposal 5: Mobility events, e.g. handover and RLF, should be considered for the evaluation of XR services.** |

# List of contributions in RAN1 #106-e

1. R1-2106457 Evaluation methodology for XR and Cloud Gaming Huawei, HiSilicon
2. R1-2106527 Remaining Issues of XR Evaluation Methodology ZTE, Sanechips
3. R1-2106630 Remaining issues on evaluation methodologies for XR vivo
4. R1-2106918 Evaluation methodology and KPIs for XR Samsung
5. R1-2106950 Evaluation methodology and performance index for XR CATT
6. R1-2107087 XR evaluation methodology FUTUREWEI
7. R1-2107280 Discussion on the XR evaluation methodology OPPO
8. R1-2107375 Remaining Issues on Evaluation Methodology for XR Qualcomm Incorporated
9. R1-2107462 Discussion on evaluation methodologies for XR LG Electronics
10. R1-2107501 On Evaluation Methodology for XR and CG MediaTek Inc.
11. R1-2107535 Remaining Issues on XR Evaluation Methodologies InterDigital, Inc.
12. R1-2107617 Evaluation Methodology for XR Intel Corporation
13. R1-2107630 Evaluation methodology for XR Ericsson
14. R1-2107656 Development of the Evaluation Methodology for XR Study Nokia, Nokia Shanghai Bell
15. R1-2107769 Considerations on XR evaluation methodology Apple
16. R1-2107887 Discussion on evaluation methodology for XR NTT DOCOMO, INC.
17. R1-2107906 Discussion on remaining issues of evaluation methodology for XR services Xiaomi