3GPP RAN 5G-ACIA Evaluations Week 2

December 14th – 18th 2020

Source: Moderator (Ericsson)

Title: Summary of discussions in Week 2

Document for: Discussion, Decision

# 1 Introduction

AT RAN#89, the following was agreed in [RP-202069](https://protect2.fireeye.com/v1/url?k=41a5db26-1f051960-41a59bbd-86fc6812c361-73f443258ff773bf&q=1&e=bc078f84-983d-45f3-ab31-19e60d911036&u=https%3A%2F%2Fwww.3gpp.org%2Fftp%2Ftsg_ran%2FTSG_RAN%2FTSGR_89e%2FDocs%2FRP-202069.zip) on providing evaluations for 5G-ACIA:

* Start an offline email-based activity to provide evaluation results for 5G-ACIA
* One company volunteers as moderator
  + Proposes a work plan to follow
  + Ericsson is willing do this
* Discussions are on the RAN1\_NR reflector
  + Email activity only during short periods (< week) distributed across the time allocated to the activity
  + No email activity in weeks before/during/after RAN1 meetings or RAN defined inactive periods
  + All companies should strive to limit email activity as much as possible
  + Outcome of the offline discussion will directly go to RAN without need for discussion in RAN1 nor need for LS from RAN1 to RAN
* Target completion by RAN#91
* At RAN#91, RAN will decide on a response LS to 5G-ACIA

The moderator made the following proposal on a timeline:

1. 12-16 October 2020
   * Discussion on which URLLC features to include in the evaluations and simulation assumptions
2. 14-18 December 2020
   * First round of simulation results
3. 22-26 February 2021
   * Second round of simulation results
4. 8-12 March 2021
   * Finalization of the report to RAN#91

During week 1, the simulation assumptions were agreed as captures in the document below:

[https://www.3gpp.org/ftp/tsg\_ran/TSG\_RAN/TSGR\_90e/Inbox/Drafts/5G-ACIA October/Agreements/Agreements week 1 5G-ACIA.docx](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_90e/Inbox/Drafts/5G-ACIA%20October/Agreements/Agreements%20week%201%205G-ACIA.docx)

For the second week, companies provided the first round of simulation results:

[https://www.3gpp.org/ftp/tsg\_ran/TSG\_RAN/TSGR\_91e/Inbox/Drafts/5G-ACIA December/Company Inputs/](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs)

The input contributions are also listed in the reference section.

This is the summary contribution for the week 2 discussions.

# 2 Company Inputs

## 2.1 Huawei/HiSilicon

[Contribution link](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/HwHiSi%20-%20Simulation%20results%20for%205G-ACIA%20in%20the%20first%20round.docx).

Other companies’ questions and comments on the results with responses:

|  |  |
| --- | --- |
| Company | Questions and comments |
| Nokia | For the MU-MIMO results, would it be possible to share more details on the assumed transmission scheme, e.g. details on the precoding, CSI acquisition and in general multiplexing of UEs?  [HW/HiSi response]:   * SRS periodicity is set as to 5 ms and also the CSI periodicity is set to 5ms. * The precoding vectors of paired UEs are adjusted according to the estimated channel vectors from SRS based on ZF. * Rank adaption is adopted but the traditional rank adaption algorithm is used. Hence for max-rank = 4, the gNB almost always selects rank=4.   Is each BS limited to 2 Tx/Rx antenna ports?  [HW/HiSi response]:   * No. The antenna setting is as in TR 38.824, i.e., 4Tx/4Rx at the BS in case of 4GHz carrier frequency.   How was the scheduling done? I.e. was SPS/CG used, or was each TB scheduled independently? If SPS/CG was used, what are the modelling assumptions for the SPS/CG scheduling? E.g. is the FDRA fixed for each UE for the entire simulation? Is the MCS selected per UE or is it the same for all the UEs? Are there any adjustments of MCS/FDRA during the simulation?  [HW/HiSi response]:   * We use dyncamic scheduling, this gives the possibility to update the MCS according to CQI. Since the data arrival is deterministic, pre-scheduling is employed as expalined in our paper. |
| Qualcomm | On “distributed MIMO”: is it mTRP Tx only, or is it ICIC and other features as well?  [HW/HiSi response]: Coherent JT transmission from all TRPs for each UE. We did not consider ICIC.  For the latency figure, are they identical for DL & UL?  [HW/HiSi response]: Yes, because we assume a symmetric frame structure and also symmetric processing times at both the BS and the UE  Please clarify how the MCS selection and radio link adaptation are used, especially in the context of “distributed MIMO”.  [HW/HiSi response]: This is done according to SRS and CSI. Similar to the single TRP multi-antenna SU/MU scheme, and the only difference is that each TRP has a power constraint. Please note that the RB allocation is performed according to CSI and SRS, but since full cell cooperation is performed, the reported values are very high and hence the highest MCS index is usually selected, i.e. 2 PRBs will usually be selected. In case there would be resources left after the allocation, the resources are then proportionally assigned to the UEs and hence a smaller MCS index could be used for transmission. |
| Ericsson | For the distributed MIMO, how was ‘the coordinated or coherent transmissions from different BSs’ done? Do the BSs coordinate to eliminate interference? Or they transmit coherently to improve SNR? Or both? But somehow the DL geometry shown in Figure 1 is worse than E/// plot, and E/// plot does not use any coordinated or coherent transmission.  [HW/HiSi response]: Coherent JT transmission is done from all TRPs to the UEs in order to improve the SINR. For the geometry, no cell cooperation is considered  For blocked or failed packets, “E2E latency is set to 1ms“. Shouldn’t the E2E latency be set to infinity or at least some value >1ms?  [HW/HiSi response]: We think it does not really matter if this value is set to exactly 1ms or to a larger value. We assume that in practice there won’t be any UE that has exactly 1ms delay. So we regard UE that <1ms latency as meeting the requirement in our simulation. But we are ok to chantge this according if it is requested.  For the number of users in Table 2 and Table 3, it’s curious how the numbers come from. They don’t seem to based on real time scheduler that allocates different amount of PRB according to actual SNR of each UE. For example, 272, 544 and 1088 are simply multiples of 272 (PRB). Does this mean that each UE gets a fixed number of 1 or 2 or 4 PRBs?  [HW/HiSi response]: The number is obtained from the simulation.   * For the SU case, the allocation is done according to the feedback CQI, but the CQI is always very high (note that we assume have full cell cooperation). As a result two RBs are allocated to each UE. * For the MU case, the allocation is done according to the CQI feedback and the UE pairing result at the BS. As a result. two RBs are allocated to each UE in case of 1088 UEs. For the 544 UEs, 4 PRBs are allocated to each UE.   Note that although we dynamically allocate the RBs to the UEs according to the feedback CQI and UE pairing results, each UE would select the highest MCS since no inter-cell interference exists (due to cell cooperation) and the feedback SINR is very high. (This would be different if interference from neighboring BSs is considered). As a result, in case of SU and MU, each UE only occupies 2 RBs initially. But if some RBs are left unused after scheduling, we will then allocate these RBs to the UEs proportionally to achieve a low SE. Hence, in case of 4-layer MU with 544 UEs, about 4RBs are finally allocated to each UE. |
| vivo | For D-MIMO, how to perform signal processing and transmission coordination? All or part of BSs can transmit/recieve by SFN way? Or they can coordiantion scheduling information, e.g.resource allocation?  [HW/HiSi response]: In D-MIMO, coherent JT transmission from all BSs is used for cell cooperation. This is similar to SFN, but in the traditional setting of SFN, each BS is equipped with one antenna port and hence, no precoding exists. Meanwhile, there is no coherent transmission in SFN and the only gain is the power and diversity gain. Also, in the traditional setting of SFN, no MU is assumed.  For dynamic scheduling, whether retransmission is enabled or not?  [HW/HiSi response]: It is not enabled. |
| Intel | Did you evaluate uncoordinated operation and how it would compare to the coordinated scheduling?  [HW/HiSi response]: He have not evaluated uncoordinated operation. Maybe we could evaluate interference from an outdoor base station?  Do you think PDCCH overhead with dynamic pre-scheduling of up to 272 UEs in every slot is accurately accounted?  [HW/HiSi response]: For PDCCH overhead, in our simulations we re-use dynamic scheduling and make the RB allocation based on the CQI feedback. But since cell-cooperation is assumed, the SINR will always be very high. Thus in practice it would be possible to use SPS/CG transmission with lower DCI overhead. Therefore, we assume DCI overhead of 20% as realistic. We can add simulations for SPS/CG in the second round if companies think it is desirable.  How the latency distribution is obtained: per packet or per UE? If it is per UE, then how a single value from all the packets of a UE is calculated?  [HW/HiSi response]: per-packet. But since the packet arrival of each UE and its transmission is fixed in each CT, the E2E latency of different packets for each UE is fixed throughout the simulations |

## 2.2 Intel

[Contribution link](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/INTEL%20-%205G-ACIA%20LS%20-%20Phase%202%20inputs.docx).

[Contribution link – updated](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/INTEL%20-%205G-ACIA%20LS%20-%20Phase%202%20inputs%20-%20updated.docx) with additional density for UL.

Other companies’ questions and comments on the results with responses:

|  |  |
| --- | --- |
| Company | Questions and comments |
| ZTE | Do you assume one baseband for all 12 BSs or separate basebands for different BSs. Is there any coordination among different BSs?  Intel: We assume deployment with 12 uncoordinated BSs |
| Nokia | How was the scheduling done? I.e. was SPS/CG used, or was each TB scheduled independently? If SPS/CG was used, what are the modelling assumptions for the SPS/CG scheduling? E.g. is the FDRA fixed for each UE for the entire simulation? Is the MCS selected per UE or is it the same for all the UEs? Are there any adjustments of MCS/FDRA during the simulation?  Intel: We use dynamic scheduling approach which could be re-interpreted as SPS/CG. The resulting behaviour is that FDRA size is fixed to the minimum allocation size of 8 PRB due to good channel quality reported, but the frequency position is explicitly randomly changing by the scheduler in order to avoid consecutive packet drops. |
| Qualcomm | What is the exact number for the % of Ues satisfying 10-4 PER for the DL simulation with 30 UE/cell?  Intel: 99.3% of UEs |
| Ericsson | In Table 1 evaluation assumptions, it has Handover margin of 1 dB. Is handover simulated? (But the agreement was ‘No explicit UE mobility (nor handovers) are modeled in the evaluations.’)  Intel: Handover is not explicitly simulated. The margin is used to randomly select among based stations with maximum RSRP within a margin, i.e. allow UE to select not the pure maximum RSRP link.  In Table 1 evaluation assumptions, was there special reason to use BS transmit power of 30 dBm? The agreement was to follow 38.824: “24 dBm per 20 MHz”, which gives 31 dBm.  Intel: We used 30 dBm from InF channel model calibration assumptions but did not change it to 31 dBm accidentally. Will use 31 dBm in the next phase. Also do not expect the results to be different due to this change.  For Figure 1(a), why was channel path gain presented? Other companies tend to show coupling loss. It’s easier for calibration if coupling loss is shown instead.  Intel: Our interpretation is that channel propagation + antenna gains etc. has total negative gain. When a definition of loss is used, the value is expected to be positive. Looking at Ericsson’s results it seems we show same statistics under different names.  For Figure 1(b), what configuration the geometry shown for? For example, BS antenna configuration is 4Tx/4Rx or 8Tx/8Rx?  Intel: It is for BS antenna 4 Tx/Rx  Regarding BLER target of 1e-5: is this a bit of overkill? With CSA=99.9999%, and survival time = 1ms, one packet error is acceptable. CSA is for two or more consecutive packet errors. BLER around 1e-3 should be adequate.  Intel: We may change this assumption in the next phase. This was mainly done assuming the consecutive error probability may not be independent. However, simulations showed that BLER target could be reduced. |
| vivo | For TDD DL-UL configuration, 1:1 DL-to-UL 7 OS DL - 7 OS UL was used. Whether DL to UL gap was considered or not?  Intel: Gap was not explicitly considered. In total from DL and UL weh ave 4 symbols overgead for control and RS and we considered that the switching gap may be absorbed into this overhead. We will check more carefully whether this is the case.  Since latency was one of the service requirement in 5G-ACIA LS, whether/how physical layer processing delay was modeled in you simulation?  Intel: We did not consider processing latency but considered all other latency components. |
| HW/HiSi | Q1: In the paper it is said that 720 UEs are simulated in total. Is the understanding correct that for 10 UEs per service area, there are 6 drops, for 20 UEs, 3 drops, for 30 UEs there are 2 drops? How about the case of 40 UEs per service area?  Intel: The text about 720 UEs was written before we added 40 UE/area into UL and DL. For this particula case, 2 drops are modelled resulting in 960 UEs.  Q2: Is it correctly understood that no cell-cooperation and no MU MIMO is assumed in the simulations?  Intel: Correct  Q3: On the number of packets that are generated to evaluate the CSA. In our view 100k packets are not sufficient to accurately estimate the CSA for these low BLERs. Not enough errors would be observed.  Intel: Agree, will consider larger number of samples next time |

## 2.3 ITRI

[Contribution link](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/ITRI_5G%20ACIA%20Simulation%20Result%20for%20InF-DH%204GHz.docx).

Other companies’ questions and comments on the results with responses:

|  |  |
| --- | --- |
| Company | Questions and comments |
| ZTE | Do you assume one baseband for all 12 BSs or separate basebands for different BSs. Is there any coordination among different BSs?  ITRI response: all 12 BS are separate and independent, there is no coordination among the BSs. |
| Nokia | What are the modelling assumptions for the SPS/CG scheduling? E.g. is the FDRA fixed for each UE for the entire simulation? Is the MCS selected per UE or is it the same for all the UEs? Are there any adjustments of MCS/FDRA during the simulation?  ITRI response: The FDRA is fixed for each UE. We use the MCS index table 2 in TS38.214, and choose the lowest MCS index to satisfy the traffic load for each UE.  The performance in terms of supported number of UEs seems significantly worse than what is reported by other companies. It was not immediately clear why this is the case, but it would be a good to understand the reasons before considering including these in the 5G-ACIA response LS. Would ITRI be able to indicate the potential reason for such low number of UEs supported?  ITRI response: When the traffic arrival just misses the pre-allocated (configured) transmission opportunity, it is postponed to the next chance, and hence the reception exceeds 1ms deadline, as shown in the following figure. |
| Qualcomm | What is the exact PER requirement?  ITRI response: We setting PER requirement 0.01% (10-4) |
| Ericsson | For section 3 simulation results table, it’s puzzling why Percentage of UEs satisfying requirements is only at the level of 70+%, while other companies‘ results for 10 UE per service area show 99+%. Some explanation text was provided about the configuration, but not very easy to understand the details.  ITRI response: When the traffic arrival just misses the pre-allocated (configured) transmission opportunity, it is postponed to the next chance, and hence the reception exceeds 1ms deadline, as shown in the following figure. |
| vivo | Both BS and UE processing delay were taken into account latency statistic or not?  ITRI response: We recently don’t take processing delay into account. |
| HW/HiSi | Q1: What MCS is used for the configured resources?  ITRI response: We use the MCS index table 2 in TS38.214, and choose the lowest MCS index that satisfies the traffic load for each UE.  Q2 In the table showing the simulation results, it is said CSA = 100%, also it is shown that slightly more than 70% of the UEs satisfy the requirements. Is it correctly understood that UE not satisfying the requirements here means that these UEs have too large latency?  ITRI response: Yes, latency of some UEs is too large, resulting in that these UEs are not satifyied.  Q3: Is the following correctly understood by us: It seems that the latency problem could be overcome with choosing a proper and more realistic traffic model (as it was discussed during the previous email discussion)? That means if the data arrival would not be random, then the configuration could be aligned with the data arrival and the latency could be reduced significantly? Prior to this simulation round we discussed the applicable traffic model. Would  Q4: For the frame structure SU with S including 12D and 2S, and the TTI of 7OS, it cannot be guaranteed that there is sufficient time to transmit and process every packet if the packet arrives randomly in the 1 ms CT. Could this be a reason for the long latency that is observed?.  ITRI response for Q3 and Q4: Yes. The main reason for the long latency is as follows. When the traffic arrival just misses the pre-allocated (configured) transmission opportunity, it is postponed to the next chance, and hence the reception exceeds 1ms deadline, as shown in the following figure.    Q5: The simulation time is 5 s, i.e., 5000 packets for each UE. We think that this is too small for the given error rates.  ITRI response: Yes it is too small for the given error rates. We will simulate more packets in the feature.  Q6: Is cell cooperation and MU is considered?  ITRI response: No, neither cooperation nor MU are considered. |
| Intel | How the latency distribution is obtained: per packet or per UE? If it is per UE, then how a single value from all the packets of a UE is calculated?  ITRI response: Per packet. |

## 2.4 Nokia

[Contribution link](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/NOKIA%20-%205G-ACIA%20First%20round%20of%20simulation%20results.zip).

Other companies’ questions and comments on the results with responses:

|  |  |
| --- | --- |
| Company | Questions and comments |
| ZTE | Do you assume one baseband for all 12 BSs or separate basebands for different BSs. Is there any coordination among different BSs?  Nokia response: all 12 BS are separate and independent, there is no coordination among the BSs. |
| Qualcomm | Please clarify the number of samples per UE (is it 2\*106)?  Nokia response: We simulate approximately 20.000 and 10.000 samples per UE. However, we simulate 100 different UE drops thus we have a large number of latency samples in total. For next round, we may slightly reduce the number of UE drops, and instead increase the number of samples per individual UE.  Regarding the statement “This is because the latency performance is impacted not only by queuing delay and interference but also by limitations to user multiplexing imposed by beamforming operation itself. )”, is the difference between DL and UL in FR 2 due to beamforming capability, i.e. fewer opportunities for scheduling UEs?  Nokia response: The same BF limitation applies for both UL and DL, but in addition, UL suffers from link budget  It would be great if the following quantities could be clarified: gNB processing delay, UE processing Delay, PUSCH preparation time  Nokia response: We assumed UE processing capability#2 for the UE latency, and used the same as minimum latency for the gNB. More specifically, the initial UE (UL) or BS (DL) processing delay and the final BS (UL) or UE (DL) processing delay are considered as N\_2/2 and N\_1/2, respectively. For HARQ retransmissions, the time interval between the initial transmission and the retransmission is the sum of N\_1+N\_2+PUCCH duration (plus frame alignment delay e.g. due to TDD frame structure). More details in our Tdoc R1-1808448. |
| Ericsson | In Appendix B simulation assumption of **FR1**,  what’s the number of UE Tx antennas and configuration? Same as Rx antennas?  Nokia response: yes, Tx antenna setup is the same as Rx antenna setup.  Any reason that BS Tx power is 27 dBm? The agreement was to follow 38.824 (24 dBm per 20 MHz), which gives 31 dBm.  Nokia response: Thanks for asking, this is a mistake. We used 31 dBm for FR1 and 26 dBm for FR2  In Appendix B simulation assumption of **FR2**, UE antenna configuration mentions “2 UE panels facing opposite directions”. Was the panel selection static?  Nokia response: The best UE panel was chosen at the beginning of the simulation and kept static. |
| Vivo | What does ‘queuing delay’ means?  Nokia response: There is data to be transmitted, but there are no PDSCH/PUSCH resources immediately available for that particular UE, thus the data is buffered until the next available opportunity. Note the we have not explicitly assumed a SPS/CG-alike scheduling scheme.  Whether BS and UE processing delay were considered in latency performance or not?  Nokia response: Yes, see answer to Qualcomm |
| HW/HiSi | Q1: According to our understanding the number of packets that are simulated in each drop is 20k in FR1 and 10k in FR2. Given the low target error rates, this seems not sufficient to produce accurate results.  Nokia response: We agree that it is not possible to provide CSA statistics per UE at the required 99.9999%-ile level. As also mentioned in our reply to Qualcomm, for next round, we will increase the number of samples per individual UE.  Q6: Is cell cooperation and MU is considered?  Nokia response: No |
| Intel | Do you think per-packet latency CDF provides sufficient understanding of the system performance? The more interest is in understanding each UE quality of service, which can rather be characterized by per-UE metrics, not necessarily latency.  Nokia response: Both ‘global’ and ‘per-UE’ metrics have been requested by 5G-ACIA. For the next round, we will simulate more samples per UE to better characterize the per-UE CSA and latency metrics. |

## 2.5 Qualcomm

Contribution links for [FR1](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/QUALCOMM-5G-ACIA_URLLC_simulation_results_1st_round_FR1.docx) and [FR2](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/Qualcomm5G-ACIA_URLLCResultsRound1_FR2.docx).

Other companies can provide questions and comments in the table below:

|  |  |
| --- | --- |
| Company | Questions and comments |
| ZTE | For FR1 evaluation, is it a correct understanding that you allocated some time domain resources dedicated for re-transmission? To allow one re-transmission, do you assume 3 symbols or 4.5 symbols for processing SPS PDSCH and preparing HARQ-ACK?  Qualcomm response: 3 symbols are needed for processing SPS PDSCH. Also, 3 symbols are needed betwen receiving a PUSCH and sending a retransmission on PDSCH. They are all compliant with TR 37.910, TS 38.2124 and TS 38211 as explained below.  Do you assume one baseband for all 12 BSs or separate basebands for different BSs. Is there any coordination among different BSs?  Qualcomm response: All 12 BS are separate and independent, and there is no coordination among the BSs except for the orthogonal retransmission phase.  In addition, it’s our understanding that assuming only 2.8 symbols for gNB processing especially for decoding PUCCH plus scheduling re-transmission is challenging.  Qualcomm response: Our calculation is based on Table 5.7.1.1.1-1 and Table 5.7.1.1.2-1 in TR 37.910 and Table 6.4-2 in TS 38.214 (Note that for 30 kHz SCS from TS 38.211 Table 4.2.1). Note that the retransmitted packet is already available in the buffer when the retransmission decision is made. Therefore, it seems practical to allow 2.8 symbols for gNB to process the PUCCH + retransmission. |
| Nokia | **FR1:**  As also commented by ZTE, UE and gNB processing times seem more optimistic than what has been assumed by other companies. For instance, 2.8 symbols are assumed from PUCCH transmission (with HARQ) to PDSCH retransmission, whereas we assume 5.5 symbols (corresponding to N2). Also, minimum DL/UL latency in the CDF (Figs. 11 and 12) is ~80 us, which also doesn’t seem very realistic.  Qualcomm response: Our calculation is based on Table 5.7.1.1.1-1 and Table 5.7.1.1.2-1 in TR 37.910 and Table 6.4-2 in TS 38.214 (Note that for 30 kHz SCS from TS 38.211 Table 4.2.1). The labels of the latency plot are indeed incorrect because the 2 SPS data symbols were not taken into account. The correct plot is attached here, where the DL and UL latencies are nearly identical:    Section 4.1: On the CSA distribution, does the reported ‘end-to-end error’ corresponds to single-error or two consecutive error case?  Qualcomm response: It’s the single-error case.  Are UEs dropped per service area or per BS? Section 4 and 5 seem to imply that there is a fixed amount of UEs in each BS.  Qualcomm response: Fixed number of UEs are dropped to each service area, while each BS may have different loads due to the pathloss association rule.  The proposed TDD configuration seems to require UE capability 5-1b of multiple DL/UL switches per slot. Should this be included in the feature list to be evaluated?  Qualcomm response: From Table 11.1.1-1 in TS 38.213, multiple DL/UL switches per slots are permitted. We think the multiple DL/UL switches per slot should be included in the feature list.  **FR2:**  Is it correct that you assume up to 4 UEs FDM? Is this opportunistic, depending on how many UEs happen to be reachable with the same beam, or is there some sort of multi-beam transmission allowing this to happen always?  Qualcomm response: Multi-beam transmission is assumed so that 4 Ues FDM always happens – provided there is traffic available.  What are the modelling assumptions for the SPS/CG scheduling? E.g. is the FDRA fixed for each UE for the entire simulation? Is the MCS selected per UE or is it the same for all the Ues? Are there any adjustments of MCS/FDRA during the simulation?  Qualcomm response: FDRA and MCS for a UE could be modified by the SPS/CG overriding symbol in each slot. More precisely, FDRA can be updated upon need for radio link adaptation. MCS is selected per UE. There are adjustments of both MCS and FDRA when the radio link adaptation indicates MCS or FDRA change-upon NACK reception. Given that not many Ues need FDRA or MCS update per slot under our SPS/CG scheme, the resources provided for SPS overriding should be sufficient. Our preliminary investigation indicated that the difference between SPS/CG and dynamic schedulings are minimal for 20 UE/cell under our SPS/CG strategy. More detailed investigation on the overhead required for effective SPS/CG is ongoing. |
| Ericsson | **FR1:**  For simulation parameters in section 2, why was the periodicity=2ms? Use case #2 has transfer interval of 1ms, which means periodicity=1ms in our understanding.  Qualcomm response: Our results continue to hold for transfer interval equal to 1 ms and the periodicity equal to 1 ms.  Could you please explain why retx BLER target is 10-4/6? How to ensure orthogonal **retx** throughout the network? No orthogonal tx if initial transmission?  Qualcomm response: The ReTx BLER 10-4/6 is found by trial and error (obviously it has to be < 10-4). The base stations in different service areas could coordinate their ReTx resources so that they could be made orthogonal throughout the network. There is no coordination among the gNBs for initial transmissions.  **FR2:**  For FR1 study, the timing was worked out to allow 1 retx, hence a high BLER target of 1e-2 could be used for initial tx. For FR2, presumably the latency requirement is easier to achieve than FR1. Isn’t it more feasible to use HARQ retx in FR2 (hence higher BLER target than 1e-4)? What’s the considerations for the setup of FR2 as compared to FR1?  Qualcomm response: Indeed, there is HARQ allowing for 1 reTx in FR2. The BLER Target for the initial transmission is 10-3. The combined BLER after the first reTx should be 10-6 considering independent transmissions  In Section 2.1.5, UE Tx power is set to 11 dBm. Any special reason to use such low value? 38.824 used 23 dBm.  Qualcomm response: Sorry for the typo. The max UE Tx Power was set to 17 dBm per antenna. |
| Vivo | What does ‘At most one retransmission occurs at any time throughout the network‘ mean? Was that mean BS have to reserve resources for retransmission at any time? Meanwhile, initial transmission can not be allocated to the reserved resources.  Qualcomm response: The orthogonal retransmission resources are allocated at the designated symbols only. The whole 100 MHz can be allocated for retransmissions at the designated symbols.  For ‘A node always transmits at full power if it is performing a retransmission‘, was it only for UL?  Qualcomm response: Both UL and DL retransmissions are performed at maximum power. The retransmissions are all orthogonal throughout the network. |
| HW/HiSi | **Q1:** In the figure 4, is the following understanding correct: If the PDSCH would be decoded correctly after the first transmission already, then the latency calculation would be terminated there, right? There is no need to wait for the PUCCH to send the ACK. Thus, in case of one-transmission, the latency would be 2.8 symbols plus PDSCH transmission time and the PDSCH decoding time, right?  Qualcomm response: Yes. A 10-symbol latency will be added for packets that need retransmisssions.  **Q2:** On Figure 4  For DL processing: It seems that the signal processing (PUCCH decoding) at the gNB receiver as well as the scheduling delay at the gNB are not considered. This combined time should be larger than the given 2.8 symbols. If we have understood correctly, then we think it is not feasible to have time-budget for 1 re-TX. Also, in our understanding, it seems that in the assumed frame structure, the gap between PDSCH to PUCCH of ACK/NACK is at most 4 symbols, this is smaller than 4.5 OS that shown in Figure 4. Are we missing here?  Qualcomm response: The original figure has some wrong labels. The revised latency figure is attached in our response to Nokia.  For UL processing: It seems that the signal processing (data decoding) at the gNB receiver as well as the scheduling delay at gNB are not considered. This time should be larger than the given 4.5OS. Similarly, the PDCCH to PUSCH time is set as 2.8 OS, smaller than 5.5OS required for UE capability #2. We don’t think that the time budget for a re-TX is feasible in the UL direction. Could you please clarify?  Qualcomm response: Please refer to our response to ZTE. In particular, the duration from PDCCH to PUSCH for ReTx is assumed to be 5.5/2 OS, which seems compliant to the current standard.  For retransmission, the decoding time at receiver for the second transmission is not considered  Qualcomm response: That’s right. Since no more transmissions are needed after the ReTx, the decoding latency could be incorporated in the upper layer. This reasoning is consistent with the latency definition in Table 5.7.1.1.1-1 in TR 37.910. The most important thing is that the packet has to arrive at upper layer before it expires.  **Q3**: Is it correctly understood that the results are not obtained directly from simulations but are calculated based on intermediate results which are obtained the simulation? Could you please clarify?  Qualcomm response: The probabilities related to retransmissions are estimated based on the statistics generated from simulation. The statistics related to initial transmissions are directly generated from simulation.  **Q4**: Is it correctly understood that the results are showing the probability that the requirements are not satisfied for all UEs, but not the percentages of UEs satisfying the requirements?  Qualcomm response: This is our main message for FR1 results: For 20 UE/cell case, 100% of UE satisfy the 1-ms latency and 10-6 error requirements. For 30 UE/cell case, not all the UEs satisfy the two requirements.  **Q5**: For FR2, we think that 100s simulation time is too short to accurately capture the results.  Qualcomm response: The PER target is 10-4, and we have 100.000 samples. Simulations with more samples can be added in the 2nd round.  **Q6**: Is cell cooperation and/or MU considered (it seems that no cooperation is adopted)?  Qualcomm response: Both are not considered, except for the orthogonal retransmission part which incurs only marginal overhead. |
| Intel | Did you evaluate uncoordinated operation and how it would compare to the coordinated scheduling?  Qualcomm response: The evaluation of uncoordinated operation will be reported in 2nd round.  Is that correct understanding, that interference is not modelled/accounted under the assumption of fully orthogonal scheduling? If that is correct, then what is potential source of different performance between different UE densities?  Qualcomm response: Yes. When there are more UEs per service area, more retransmission resources are required to maintain 10-6 end-to-end error for each UE because on average there are more UEs that require retransmissions. |

## 2.6 vivo

[Contribution link](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/vivo-5G-ACIA%201st%20round%20URLLC%20evaluation%20results.DOCX).

Other companies’ questions and comments on the results with responses:

|  |  |
| --- | --- |
| Company | Questions and comments |
| ZTE | Do you assume one baseband for all 12 BSs or separate basebands for different BSs. Is there any coordination among different BSs?  vivo response: All 12 BSs are independently scheduled, and there is no coordination among them. |
| Nokia | How was the scheduling done? I.e. was SPS/CG used, or was each TB scheduled independently? If SPS/CG was used, what are the modelling assumptions for the SPS/CG scheduling? E.g. is the FDRA fixed for each UE for the entire simulation? Is the MCS selected per UE or is it the same for all the UEs? Are there any adjustments of MCS/FDRA during the simulation?  vivo response: SPS/CG was used in our simulation, and the period is 1ms. FDRA and MCS selection are based on CSI reporting. |
| Qualcomm | What is the CSA assumption?  vivo response: The CSA assumption is 99.9999%, and the metric is referred TS 22.104. |
| HW/HiSi | **Q1:** Table one says “satisfying the requirements” What requirements are meant, is it latency, CSA or both om them? Could you please clarify?  vivo response: When we call one UE satisfying the requirements, it means CSA≥99.9999% and latency < 1ms.  **Q2**: IS cell cooperation and/or MU considered (it seems that no cooperation is adopted)?  vivo response: All BSs are independently scheduled. We did not adopt cooperation and MU transmission. |
| Intel | How the latency distribution is obtained: per packet or per UE? If it is per UE, then how a single value from all the packets of a UE is calculated?  vivo response: The latency distribution is obtained per UE. For a given UE, a single latency value is from the average of all the packets latency of the UE. |

## 2.7 Ericsson

[Contribution link](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/Ericsson%205G-ACIA%20Simulation%20Results%20Round1.zip).

Other companies’ questions and comments on the results with responses:

|  |  |
| --- | --- |
| Company | Questions and comments |
| ZTE | Do you assume one baseband for all 12 BSs or separate basebands for different BSs. Is there any coordination among different BSs?  [Ericsson] No coordination between BSs. All 12 BS are separate cells and have separate baseband modules. |
| Nokia | What are the modelling assumptions for the SPS/CG scheduling? E.g. is the FDRA fixed for each UE for the entire simulation? Is the MCS selected per UE or is it the same for all the UEs? Are there any adjustments of MCS/FDRA during the simulation?  [Ericsson] FDRA is fixed for each UE for the entire simulation. The MCS is selected per UE, and selected once in the beginning of the simulation. No adjustments of MCS/FDRA during the simulation.  Minimum latency is ~100 us, which is a bit too low for 2 OS TTI and 30 kHz SCS. Are realistic processing times taken into account?  [Ericsson] Thanks for pointing out that the minimum latency of ~100us is a bit too low. We will double check our latency modeling in the simulator. Also, if/when companies converge on a common way to measure E2E latency (i.e., the dedicated email thread), we will update our simulator to reflect that. |
| Qualcomm | Please elaborate the sentence “Since packet arrival is known by gNB, allocation in time and periodicity is optimized so that the alignment delay is minimized.” What quantities are optimized to minimize alignment delay?  [Ericsson] Mini-slot for CG/SPS resource allocation is chosen to match the timing of packet arrival. Additionally, the CG/SPS periodicity equals the transfer interval (packet periodicity). |
| vivo | How to define the maximum supported UEs per service area? Define as the maximum number of UEs per service area achieve 99.9999% CSA for more than 99% of users?  [Ericsson] Not sure about the meaning of the question. In our simulation, the number (e.g., 10) of UEs per service area are deployed. The packets are scheduled and transmitted for each UE. Then the statistics are collected for CSA and delay for all UEs in the factory hall. |
| HW/HiSi | Is cell cooperation and/or MU MIMO adopted?  [Ericsson] No, no cell cooperation or MU-MIMO. SU-MIMO is used. |
| Intel | How to interpret the latency CDF point with ‘0 ms’ value?  [Ericsson] There is no 0 ms measures. Latency measures are collected in bins with 0.01ms step. Flat line in the beginning just an artifact from plot script.  How the latency distribution is obtained: per packet or per UE? If it is per UE, then how a single value from all the packets of a UE is calculated?  [Ericsson] Latency for all packets from all UEs in UL and DL are summarized in one CDF. |

## 2.8 ZTE

[Contribution link](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/ZTE-5G-ACIA%20evaluations%20-%201st%20round%20of%20simulation%20results.docx).

Other companies’ questions and comments on the results with responses:

|  |  |
| --- | --- |
| Company | Questions and comments |
| Nokia | What are the modelling assumptions for the SPS/CG scheduling? E.g. is the FDRA fixed for each UE for the entire simulation? Is the MCS selected per UE or is it the same for all the UEs? Are there any adjustments done to the MCS/FDRA during the simulation?  ZTE: FDRA is not fixed. We first assume a pre-defined FDRA and MCS, while it may or may not be adjusted based on the CSI feedback or SRS measurement. MCS is selected per UE.  Minimum latency is ~250 us, which is a bit too low for 5 OS TTI and 30 kHZ SCS. Are realistic processing times taken into account?  ZTE: For now, we assumed only 2 OS, i.e., floor(N1/2) or floor (N2/2) for 30kHz SCS, for processing initial transmission for periodic traffic scheduling. In our understanding, processing transmission may require less time compared to processing reception. But, we may consider other processing times depending on further discussion this week. |
| Qualcomm | How do the base stations coordinate together? What are the technologies involved, such as mTRP, ICIC or other features as well?  ZTE: In our simulation, we assumed a same scheduler for resource allocation for all BSs. The intention is to cancel or mitigate the ICI. |
| Ericsson | For latency figures (Fig 3 and 4), why are the CDF curves in staircase shape? Is it related to packet arrival being generated with symbol granularity?  ZTE: Yes, as explained in our paper, the packet arrival is randomly generated in symbol level in our simulation for now. We may consider more accurate data arrival in later phase.  For Table 1 RU results, it is puzzling why RU is so low. Our back of envelope estimate is, using RU=10.08% for 50 UEs per service area as in ZTE Table 1, most UEs are allocated with 1 PRB. This seems very low. For example, as a reference point, Intel’s RU results (Table 2 and Table 3) are approximately 4 times as high as ZTE’s for both DL and UL.  ZTE: The actual allocated number of RBs is based on the CSI feedback/SRS measurement. When the number of UEs is small, e.g., 10 UEs per service area, different UEs could be FDMed assuming there is coordination among UEs. In such case, the SNR is quite high since there is no interference. We find the Rank could reach to 4 for DL and 2 for UL for almost all UEs. So, one RB is actually allocated for most UEs with very high MCS. In case the number of UE is in the factory is larger than 273, e.g., 50 UEs per service area, UEs cannot be always FDMed and there will always some interference. We assumed some coordination among BSs to reduce the interference, and find allocating more RBs which will incur very high interference would not help much. Based on our current scheduling, we still find most UEs used 1 RB in most of transmissions. |
| vivo | Whether BS and UE processing delay were considered in latency performance or not?  ZTE: 2-OS BS/UE processing delay is assumed for initial transmission for now, we may consider other values in later phase evaluation. Note that, we didn’t enable re-transmission, so the processing time at receiver side is not considered in E2E latency. |
| HW/HiSi | Q1: Which MCS table is used? Is it correctly understood that a packet only occupies one RB, and hence the 256QAM table is used and the highest MCS index is used?  ZTE: qam64LowSE table is used. The actual allocated number of RBs is based on the CSI feedback/SRS measurement, we find for most of UEs in most of transmissions, one RB is allocated.  Q2: Since cell cooperation is adopted, why is the RU shown in Table 1 is so low? For 40 UEs per area, about 40\*12/2 = 240 RBs should be used in each TTI in the factory. Hence the RU should be 240/273? Is there something that we are missing?  ZTE: We didn’t assume SFN-like scheduling. In our evaluation, all BSs has its own physical resources, i.e., 273 RBs, for transmission. But we assumed some coordination among BSs to mitigate the interference. In addition, based on our frame structure, there could be two transmission occasions within 1 ms latency. So, in your example, if one RB is used for each UE in each service area, the RU would be (40\* 12)/ (273\*2\*12) = 7.33%. You can find more details in response to Ericsson above.  By the way, we find there is a typo for the RU of UL for 40 UEs, it should be 7.41% instead of 6.07%. The reason is that, we intended to simulate 5e5 TTIs, while we stopped this simulation at ~4.1e5 TTIs for timely submission of the results. When calculating the RU, it is still averaged over 5e5 TTIs. But, there is no such problem for other cases.  Q3: It seems that SU MIMO is adopted in the simulation, is this a correct understanding? If yes, then for the case of of 50 UEs per service area (i.e., 600 UEs in the factory), the resources would be overloaded, then how is it decided which UE would be served?  ZTE: We didn’t assume SFN-like scheduling. In our evaluation, all BSs has its own physical resources, i.e., 273 RBs, for transmission. Theoretically, it could serve a maximum of 273 (RB)\*2 (TO) UEs per service area for SU MIMO case. |
| Intel | What number of samples was used to conclude that CSA is 100% for certain scenarios?  ZTE: In our initial simulation, we used 5\*e5 TTIs(slots), i.e., 2.5e5 packets per UE. More TTIs may be used in the next round of simulation. |

# 3 Updates of simulations assumptions and missing simulations

## 3.1 Company input review phase

In the table below, companies provided inputs on need for changes in simulation assumptions and what additional simulations that should be performed for the second round of simulations.

|  |  |
| --- | --- |
| Company | Input |
| Nokia | The antenna assumption for FR2 may be too restrictive, with only 2 Tx/Rx antenna ports in the BS. When analogue BF is used, this limits the multiplexing capability and leads to starvation of slots to transmit while a lot of PRBs are unused. We should consider an antenna system that allows for larger number of Ues to be multiplexed to transmit/receive at the same time.  Not exactly related to further simulation needs, but realizing that some companies are simulating a single gNB with 12 RRUs having some sort of joint scheduler, while others are simulating 12 independent gNBs without coordination it would seem important to categorize the results in the final output so that it is possible to understand that different network setups lead to different performance. |
| Qualcomm | Our proposed scheme assumes the following: CSA metric with no consecutive errors; gNBs do not coordinate except for orthogonalized retransmissions, which is a special form of mTRP; Multiple uplink/downlink switchings in a slot; Processing delays compliant with TR 37.910, TS 38.2124 and TS 38211 are used for our HARQ strategy.  Most companies use the same CSA metric where no consecutive errors are permitted. It seems appropriate to make this CSA assumption mandatory in the 2nd round.  Also, the set of permitted coordination strategies among gNBs should be made more specific. The companies that assume gNB coordination should elaborate the coordination strategies being used because different coordination strategies have different processing requirements. |
| Ericsson | There is a need to calibrate the simulator among companies. We noticed at least two sets of DL geometry curves for 4GHz:   1. E/// and QC have very similar DL geometry curves; 2. HW and Intel have very similar DL geometry curves;   But (b) are significantly worse than (a), for example, about 4 dB worse at CDF=50% and 80%. It would be good to align this basic setup first. |
| Vivo | We need to align physical layer processing delay modeling. In our understanding, in addition to alignment delay and transmission duration, BS and UE processing delay also should be taken into account. Otherwise, the simulation results cannot reflect real deployment.  Since different companies have different understanding on the maximum supported Ues per service area, we need to further clarified. |
| HW/HiSi | **E2E latency**: One critical factor is how we define the E2E latency. One company reports that there is enough time for one re-TX whereas all other companies think that there is only time for a one-shot. But even among those companies, there seems to be a different understanding what is factors are contributing to the E2E latency. Maybe we should spend some more effort on this to align our view?  **Cell cooperation:** Two companies assumed cell cooperation whereas all other companies did not assume cell cooperation. For the given deployment scenario, it seems that cell cooperation can be easily achieved and the performance is very good and practicable since interference can be avoided. We would encourage more companies to study cell-cooperation.  **MU-MIMO**: Only one company studied MU-MIMO. For the given scenario, it is a very good feature to further increase the performance. It would be nice if more results also could be presented. Maybe it would be possible to define some 3 cases for simulations and to compare the results?   * No cell cooperation, SU MIMO * Cell cooperation, SU MIMO * Cell cooperation, MU-MIMO   **Calibration:** Intel, ITRI, E///, QC and HW/HiSi present results (mainly coupling loss and geometry) for calibration. There are some small differences in the results. Also, it seems that there are small differences in the simulation assumptions of different companies. Maybe we could try to prepare a table in order to compare the detailed simulation parameters from different companies, and then generate the calibration figure to compare the calibration results?  **Number of samples that should be generated:** It seems that many results are based on few samples. It would good to give a target for the produced number of packets depending on the error rate, e.g. 100/error\_rate? |
| Intel | Agree with Ericsson that it would be great to have more alignment on basic geometry curves.  Clearer categorization between coordinated and uncoordinated scheduling results is desirable.  We can now narrow down to InF-DH explicitly. Also, can confirm options of traffic arrival for DL and UL, since no company proposes to revise it in this phase.  May be good to agree on a minimum number of packets per UE and minimum number of UEs / network drops modelled for more confidence in results. |
| ITRI | **Calibration:** Agree with Ericsson and HW/HiSi.  There is a need to calibrate the simulator among companies, and prepare a table in order to compare the detailed simulation parameters. |
| Nokia | **Calibration:** Should be OK to do geometry calibration. |

## 3.2 Follow-up after review phase

### 3.2.1 Calibration

Several companies mentioned the need to calibrate the simulations from different companies based on the geometry curves being different. An Excel sheet will be provided for including the simulation parameters. Companies can add additional parameters if seen needed for the calibration exercise. The Excel sheet is found in this folder.

Also, companies not providing results for calibration are encouraged to do so.

**Conclusion:**

* The final Excel sheet can be found here.

### 3.2.2 Antenna assumptions for FR

Nokia antenna assumption for FR2 with 2RX/TX may be too restrictive, limiting the multiplexing capabilities. In the table below, companies can give input on this and also indication what antenna assumptions should be used for the additional simulations.

|  |  |
| --- | --- |
| Company | Input |
| ZTE | We are open to consider additional assumptions for FR2, e.g., 4RX/TX with the same antenna element configuration as TS 38.824. But we prefer to make it an optional assumption. |
| ITRI | We support antenna assumption with 2 RX/TX as mandatory.  We also encourage companies to perform other antenna assumption. |
| Nokia | We would not seek to change the assumptions made in October, but allow additional flexibility if the proponent sees that the selected antenna assumption severely restricts the system’s ability to utilize PRBs. |
| Ericsson | In our view, typical BS implementation is 2Tx/2Rx for FR2. It is preferable to keep 2Tx/2Rx as mandatory BS antenna configuration for FR2. We are also open to allow higher number of Tx/Rx as optional for FR2. |

**Conclusion:**

* 2RX/TX is still the baseline
* Results for additional configurations can be provided

### 3.2.3 Cell coordination

The use of cell coordination varies among companies. Companies should describe the level of cell coordination in the simulation assumption Excel sheet.

For the final report, it is probably good to separate the results with and without coordination. To see the interest in providing more results with cell coordination in round 2, companies can indicate this below.

|  |  |
| --- | --- |
| Company | Input |
| HW/HiSi | We think that cell coordination is very easy to realize in the target deployment scenario. Therefore, we would like to encourage more companies to perform these simulations. |
| ZTE | In our view, there could be at least three cell deployments.  Alt 1: All 12 BSs are separate and independent, and there is no coordination among the BSs.  Alt 2: SFN-like operation, i.e., 12 BSs share the same BBU so as to perform full coordination among BSs.  Alt 3: 12 BSs have its own physical resources while enjoy the same scheduler to enable full coordination among BSs.  For now, we assumed Alt3. We think it would be a good exercise to compare the performance of different alternatives, especially when the number of UEs is large and the resources is overloaded. |
| Intel | Prefer not to mandate cell coordination, but it is fine to encourage companies to perform additional studies and separate the results in round 2. |
| ITRI | The coordination methods would be diversified among companies, and it may be hard to converge the simulation result. We support no cell coordination as mandatory. |
| Nokia | In our view the baseline setup should be no coordination (Alt 1 of ZTE list), but it is interesting to consider full coordination as well, especially if there is an obvious system bottle neck that can be mitigated/eliminated with coordination (Alt 3, which we understand was assumed at least by HW and ZTE). |
| Ericsson | No coordination among BSs should be mandatory.  Alt 3 in ZTE list can be included as optional.  Alt 2 in ZTE list does not improve the system capacity to support more UEs, hence can be eliminated. Also, so far no companies simulated Alt 2. |

**Conclusion:**

* No coordination is baseline
* Results with cell coordination can be provided

### 3.2.4 MU-MIMO

Huawei stated that MU-MIMO should be used. To see the interest in providing more results with MU-MIMO, companies can indicate if they will provide this for round two.

|  |  |
| --- | --- |
| Company | Input |
| HW/HiSi | MU-MIMO can further increase the capacity. We would be willing to provide more results if assumptions are going to be updated in this round. |
| ZTE | We may consider to use MU-MIMO if SFN-like operation is considered. |
| Intel | Will strive to get MU-MIMO results, but prefer this to be optional. |
| ITRI | We support SU-MIMO as mandatory. |
| Nokia | MU-MIMO or spatial multiplexing with analogue beams (FR2) could be considered for better user multiplexing at least in FR2. We could also leave digital domain MU-MIMO out and list it as a potential performance enhancing candidate. In FR1 and 273 PRBs it is not immediately obvious if MU-MIMO is needed over FDM though. |
| Ericsson | Only SU-MIMO as mandatory. |

**Conclusion:**

* SU-MIMO is baseline
* Results with MU-MIMO can be provided

### 3.2.5 Latency

Vivo and Huawei indicated the need to further discussion the latency on whether to simulate with one-shot and/or HARQ retransmissions and how to calculate the latency. To have a more interactive discussion on this, a separate discussion on the RAN1 NR reflector was done.

#### 3.2.5.1 Company inputs

Below is a summary of the company inputs.

Q1: Given the tight latency budget of 1ms, which transmission option is preferred?

* Option 1: one-shot transmission
* Option 2: with HARQ re-TX

|  |  |
| --- | --- |
| Company | Input |
| HW/HiSi | Option 1.  According to our analysis in Q2 for the E2E latency Option 2 is not feasible. The value for one re-transmission is at least T =  3OS + 2OS + 5OS+1OS+14OS+2OS+3OS = 30OS. Note that this is for the case that no alignment delay is considered.  Considering the random arrival of packet, the alignment delay can be additional 6OS depedning on the frame staructure. Then the value is at least increased to 36OS. |
| Qualcomm | Both options.  The potential of HARQ should be explored because the timeline seems possible if gNB processing power is sufficiently high and if the gNBs need not coordinate to perform complex scheduling such as coherent joint transmission.  Given that the probability of having > 30 packets throughout the network is much lower than 10-6, for the 30 UE case, 3OS could be sufficient for a base station to decode PUCCH and then schedule DL retransmissions. Note that the retransmission packet could be prepared before receiving the PUCCH. Similar reasoning holds for scheduling delay of uplink retransmissions.  [HW/HiSi]: Is this the following sentence from your comment really feasible “*Note that the retransmission packet could be prepared before receiving the PUCCH*”?  I think it is not feasible. The reason is that before decoding the PUCCH, the gNB does not know which UE has failed to successfully receive the initial PDSCH. Therefore, if in advance operation would be done, the gNB would need to prepare the retransmission of all UEs, not just for the failed ones. This computation would occupy the CPU in other TTIs, and would significantly reduce the available computational resources to perform the other signal processing that is necessary during that TTI. Please keep also in mind that there is a very large number of UEs that the gNB has to deal with, e.g., 480 in total and 40 per BS even without cell cooperation. The BS already needs to prepare the initial transmission of DL, the initial transmission of UL, all PUCCHs and all PUSCHs decoding, and the scheduling/preparation of some UEs’ retransmission within 1ms.  Qualcomm response: Under the CSA assumption of packet zero survival < 1 ms, the simulation results from us and other companies indicate that the capacity would be at most 30 without gNB coordination. Since most packets are SPS and the FDRA for most packets in a slot remain unchanged compared to the slot 1ms ago,  gNB could indeed prepare all packets in advance and modify the FDRA or MCS if necessary. If gNB is given sufficient memory and computing power (e.g., commercial GPUs could be used), the gNB scheduling and processing delay should not be a bottleneck to air interface latency.  Given that the probability of seeing > 30 retransmissions throughout the network is negligible and given that the codebook for a retransmitted packet could be determined (or pre-configured) at the time when the packet arrives, the processing delay for retransmission could be made much shorter than that for initial transmissions.  The achievability of latency timing for HARQ process is implementation dependent. It is analogous to the achievability of coherent joint transmissions (JT), which is also implementation dependent. While coherent JT is considered feasible given sufficient resources, it seems reasonable to state that HARQ is feasible given sufficient resources.  HARQ should be considered for FR2. Delay budget analysis according to TR 37.910 allows for 1 retransmission within 1 ms, for PDSCH & PUSCH mini-slot configurations of 2 and 7 symbols mini slot. With specific setup, 1 retransmission is possible for a 14 symbol PDSCH & PUSCH transmission.  In general, certain gNB and UE configurations and implementations, allow for 1 retransmission even in FR 1. There is no requirement for agreement at this stage on the exact gNB configuration, the most important being that the 1 ms latency can be met under some configurations.  Hence, the feature of HARQ should not be excluded. (Agreement with Nokia.) |
| vivo | Option1  Based on our evaluation assumptions as presented in Q2, it is impossible to perform one time retransmission. But it may be possible only for the best case (no alignment delay and shorter Tx duration, such as 2 or 4 symbols) |
| Nokia | It would look to us that there is typically no time for a Re-Tx, and that’s why 1-shot is what we have been assuming. However, if with some setups there is time for a Re-Tx, that is a feature that can definitely be used. There doesn’t seem to be a strict need to agree to a given way of operating the system as long as the latency bound is kept. |
| ZTE | Option 1 for FR1, Option 2 for FR2.  More details please find our answer in Q2 |
| Ericsson | We do not see the need to impose a rigid limitation to use Option 1 vs Option 2.  In principle, if there is sufficient time to have HARQ retransmission, it is preferably to use HARQ retransmission since this improves system performance.  How the timeline works out depends on various factors that are up to each company’s choice. For example, TDD DL/UL pattern, dynamic scheduling vs DL-SPS/UL-CG, FR1 vs FR2. Thus, for companies that the timeline works out, HARQ retx should be allowed. For companies that the timeline does not work out, it’s natural to rely on initial transmission only.  So far, at least for FR1, most companies’ timeline does not work out to enable HARQ retx. Thus initial-transmission-only can be captured as the main results for FR1 (possibly with additional results included for one-HARQ-retx and timeline description). |
| ITRI | ITRI is fine with two options. We don't have special concerns. |

Q2: Which components are in your view included in the E2E latency for DL and for UL? Do you think that tables 5.7.1.1.1.-1 for DL and 5.7.1.1.2.-1 for UL from TR 37.910 should be used, or are any additional components needed as well? What values do you think are reasonable to assume for the different components?

|  |  |
| --- | --- |
| Company | Input |
| HW/HiSi | We would be fine with using the tables as starting point, but think one additional component could also be taken into account (e.g. T9) in our description below. But we also think that some of the values assumed for some components in the tables are not really achievable in practice (e.g. 2.1 and 2.4) and could be further discussed.   For DL, the E2E delay is from the time instance that the packet arrives at the PDCP layer at the BS to the time instance that it is successfully decoded and delivered to the PDCP layer at the UE. Then, the E2E latency includes the following components:     1. **The preparation time T1 at the gNB**, i.e., the time from the data arrives at the PDCP to the time that the PDSCH is ready. This value T1 depends on the transmission strategy. For example, if dynamic scheduling is used, T1 would be large since the scheduling delay and the preparation of PDCCH would be considered. If SPS PDSCH or pre-scheduling is used, only the time for TB assembling and PHY processing of the PDSCH are needed. For the latter case, we think T1 = 3OS is reasonable. 2. **The alignment delay T2 at the gNB**, i.e. the gap between the time instance that the gNB finishes the preparation of the PDSCH and the starting time of the earliest available DL TTI. The value of T2 may be random. The minimum value is 0, and the maximum value of T2 depends on the frame structure and the TTI duration, maybe it is equal to the DL-UL switching periodicity in some cases. 3. **The transmission time T3 of PDSCH,** depending on the TTI duration. For example, T3 is 2OS or 6OS. 4. In case of one transmission, 5. **The decoding of the PDSCH and the processing time to send the data to the PDCP layer at UE,** denoted as T4. We think T4 = 3OS is reasonable. 6. In case of one re-transmission, 7. **The decoding of PDSCH and preparation of PUCCH carrying ACK/NACK**, T5, depending on the UE capability. For SPS or pre-scheduling, T5=N1 = 4.5OS can be assumed. 8. **The alignment delay T6 for PUCCH transmission**. Since the PUCCH can be transmitted on any UL symbols, the delay is the time that the gap between the time instance that the UE finishes the preparation of PUCCH and the starting time of UL symbols, at least 0.5 OS. 9. **The transmission time of PUCCH T7**, which can be set as 1OS or 2OS. 10. **The decoding time, scheduling delay and the preparation time of PDCCH and PDSCH at the gNB**, denoted as T8\_1, T8\_2 and T8\_3 respectively. Note that the scheduling delay is very large, the total time T8 = T8\_1+T8\_2+T8\_3 would be also large, at least 0.5ms, i.e., 14 OS in our understanding. 11. **The alignment delay at the gNB** for retransmission T9 at BS. 12. **The transmission time of PDCCH and PDSCH T10**. For brevity, assume T10 = T3. 13. **The decoding time of the retransmission**, which is still assumed as T4 for brevity.   Then the E2E latency is then calculated to   1. for one transmission is T = T1+T2+T3+T4, 2. for one retransmission is T = T1+T2+T3+T5+T6+T7+T8+T9+T10+T4 = T1+T2+2\*T3+T4+T5+T6+T7+T8+T9.   For UL, the E2E delay is measured from the time instance that the packet arrives the PDCP layer at the UE to the time instance that it is successfully decoded and delivered to the PDCP layer at the BS. Then, the E2E latency includes the following parts:   1. **The preparation time S1 at the UE assuming** CG for brevity. We think S1 = 3OS is reasonable. 2. **The alignment delay S2 at the UE**, i.e., the gap between the time instance that the UE 3. The transmission time S3 of CG PUSCH, depending on the TTI duration. For example, S3 is 2OS or 6OS. 4. In case of one transmission, 5. **The decoding of PUSCH and processing time to send the data to the PDCP layer at the BS,** denoted as S4. We think S4 = 3OS is reasonable in our view. 6. In case of one re-transmission, 7. **The decoding of the PUSCH**, **the scheduling delay and preparation of UL grant, S5**, this is depending on the UE capability. S5 is often large due to the scheduling delay. 8. **The alignment delay S6** for UL grant. 9. **The transmission time of PDCCH carrying UL grant** S7, which can be set as 1OS or 2OS. 10. **The decoding time of UL grant and preparation of PUSCH** S8. For brevity, S8=N2=5.5OS is assumed. 11. **The alignment delay at the UE for retransmission PUSCH** S9 at BS. 12. **The transmission time of PUSCH** S10. For brevity, assume S10 = S3. 13. **The decoding time for retransmission**, which is still assumed as S4 for brevity.   Then the E2E latency in UL is then:   1. for one transmission is S = S1+S2+S3+S4, and the E2E latency   including one retransmission is S = S1+S2+S3+S5+S6+S7+S8+S9+S10+S4 = S1+S2+2\*S3+S4+S5+S6+S7+S8+S9. |
| Qualcomm | The two tables are used for prior simulation studies, and the timeline outlined therein should be achievable given enough computing power.  Even with current technology, we disagree with Huawei that the scheduling delays need span more than 10 symbols. Here is the reasoning: Suppose the initial transmissions (either UL or DL) fail *independently* with BLER target ρbler≤10-2. Also, define the probability of ReTx overflow as  where ND1andND2 denote the total number of UEs and the number of retransmissions needed respectively. It can be verified by simulation tools that for 30 UE/cell (360 UEs in total), the probability of having more than 30 retransmissions P(ρbler, ND1,ND2) is less than 10-4/6, which has the same order of retransmission BLER target. In other words, all the gNBs need to schedule fewer than 30 retransmissions in total, which is not a difficult task to complete.  In addition, simple or complex retransmission scheduling could be preconfigured for each failure pattern, thus further minimizing the scheduling latency.  Therefore, we believe that our presented frame structure for HARQ process is practically achievable even with existing technologies. If many companies think that coordination among gNBs for retransmission introduces too much latency, we are open to refine the retransmission strategy so that the gNBs can act independently.  Based on tables 5.7.1.1.1-1 and 5.7.1.1-2.1 from TR 37.910, 1 HARQ retransmission is possible in FR 2. |
| vivo | Prefer to reuse tables 5.7.1.1.1.-1 for DL and table 5.7.1.1.2.-1 for UL from TR 37.910 as starting point for E2E latency analysis. The components of DL/UL data transfer time are all taken into account for E2E latency statistic.   The following evaluation assumptions are considered for E2E latency analysis.   1. TDD DL-UL configuration 6D:2S:6U 2. PUSCH and PDSCH processing capability 2   For DL E2E latency, it can be observed from the below table,   1. At least 11OS is required for one-shot transmission. 2. At least 28OS is required if one time retransmission is enabled. Considering frame alignment delay, E2E latency would beyond latency requirement 1ms.   Note that [X, Y] in the table means the time needed for the best case and worst case.  Table 1. DL user plane latency   |  |  |  | | --- | --- | --- | | ID | Component | Value(OS) | | 1 | DL data transfer | [11, 19] | | 1.1 | BS processing delay | 5.5/2 | | 1.2 | DL Frame alignment | [0, 8] | | 1.3 | TTI for DL data packet transmission | 6 | | 1.4 | UE processing delay | 4.5/2 | | 2 | HARQ retransmission | [17, 33] | | 2.1 | UE processing delay | 4.5/2 | | 2.2 | UL frame alignment | [0, 8] | | 2.3 | TTI for ACK/NACK transmission | 1 | | 2.4 | BS processing delay | 5.5/2 | | 2.5 | Repeat DL data transfer from 1.1 to 1.4 | [11, 19] | |  | Total one retransmission user plane latency for DL | [28, 52] |   For UL E2E latency, same observations can be obtained as DL. |
| Nokia | We would prefer using the 37.910 tables as the results of this work will very likely be compared to the IMT-2020 self evaluation and any delta is assumed to originate from the Rel-16 features or the 5G-ACIA scenarios. Hence feel it may not be good to tweak the processing time assumptions and thus it is somewhat fruitless to debate on whether or on what sort of a platform those are achievable. We have already deemed them as valid and should stick to that. |
| ZTE | We are fine with using the tables in TR 37.910 for latency calculation, with adding the alignment delay at gNB side for scheduling re-transmission as suggested by Huawei. This also aligns the discussion in Rel-16 URLLC SI per our understanding.  Regarding the alignment delay and transmission duration, it depends on the companies' reported frame structure and signal overhead (PDCCH/PUCCH/SRS). In our evaluation, there are 6-OS for DL/UL per slot, and one symbol is assumed for PDCCH and one symbol for PUCCH/SRS, resulting in 5-OS TTI for scheduling DL SPS PDSCH/CG PUSCH.  Regarding the processing times, if we'd like to explicitly agree on some values, we prefer to use the values in R1-1901472 as agreed in Rel-16 URLLC SI. The related parts are copied in the end of my reply for easy reference.  For FR1, we agree with Huawei that there is no time for one re-transmission. For FR2, if we use processing times in R1-1901472, the E2E latency for one re-transmission without considering alignment delay would be:   * gNB processing time T1 = N2/2+X = N1/2+X = 18 OS. * PDSCH duration: T3 = 2~6 OS. * Decoding PDSCH and preparing PUCCH: T5 = N1 = 20 OS. * PUCCH duration: T7 = 1 or 2 OS * PUCCH-to-PDCCH processing time for re-transmission of the PDSCH: T8 = N2+X = N1+X = 28 OS. * PDSCH duration: T3 = 2~6 OS. * Decoding PDSCH: T4 = N1/2 = 10 OS   In total, it is 81~90 OS. The total OS available in case of 120 kHz SCS is 8\*14 = 112 OS. Depending on the frame structure and data arrival, we think it is possible to schedule one re-transmission in FR2.  *gNB’s processing time for transmission of the initial PDSCH and gNB’s PUCCH-to-PDCCH processing time for re-trasnmission of the PDSCH:*  *· Case1: UE’s N2/2 + X for scheduling the initial PDSCH and UE’s N2 + X for re-transmission.*  *· X = 2/4/8 symbols for SCS = 30/60/120KHz, respectively.*  *· For the case of grant-free PUSCH, the latency of the initial transmission must also include the UE’s processing time given as UE’s N2/2*  *· gNB’s PUSCH-to-PDCCH processing time (note that PDCCH alignment has to be included separately) is UE’s N1 + X*  *· X = 2/4/8 symbols for SCS = 30/60/120KHz, respectively.*  *·  gNB’s decoding time for the last PUSCH is UE’s N1/2 + X*  *· X = 2/4/8 symbols for SCS = 30/60/120KHz, respectively.*  *· For the DL study, it is assumed that N2=N1 when calculating gNB processing time. This assumption applies only to the Rel. 16 based analysis.*  *· For the UL study, it is assumed that N2=N1 when calculating gNB processing time. This assumption applies only to the Rel. 16 based analysis* |
| Ericsson | Prefer to reuse tables 5.7.1.1.1.-1 for DL and table 5.7.1.1.2.-1 for UL from TR 37.910, including the values for the components.  Some components don’t have a predefined value, e.g., TDD UL/DL pattern, transmission duration. These are up to each company’s choice.  Regarding the values for gNB and UE processing time, we do notice that TS 38.824 appendix A.4 provide alternative values, which are quoted by ZTE. However, to keep everything consistent, we suggest to follow the component values in tables 5.7.1.1.1.-1 and table 5.7.1.1.2.-1 of TR 37.910.  Thus, gNB processing time (or BS processing delay) is Tproc,2/2 ( from 37.910), not N2/2 + X  (38.824 appendix A.4). |
| ITRI | ITRI is fine with using the tables in TR 37.910 for latency calculation. We don't have special concerns. |

#### 3.2.5.2 Summary and proposals

Based on the discussion, the following summary was made including proposals.

**Summary on whether to simulate with one-shot and/or HARQ re-TX**

According to companies’ feedback, whether they prefer to perform simulations for one-shot and/or with HARQ re-TX, following feedback has been received:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Company | FR1 | | FR2 | |
| One-shot | With re-TX | One-shot | With re-TX |
| HW/HiSi | Yes |  | yes |  |
| Qualcomm | Yes | Yes | Yes | Yes |
| Vivo | Yes |  |  |  |
| Nokia | Yes | Open |  |  |
| ZTE | Yes |  |  | Yes |
| Ericsson | Yes | Open |  |  |
| ITRI | Open | Open |  |  |
| Intel | Yes | Open | Yes | Yes |

For FR1 all companies are supportive of simulations with one-shot transmission, but there is no consensus to preclude Option 2 either. For FR2 not much feedback has been received. It seems not feasible to encourage or to preclude any option.

In order to achieve more comparable results companies are encouraged to provide simulations with one-shot for FR1. If companies want to provide additional simulations with re-TX for FR1, then this should not be precluded. But it should become clear how the latency has been calculated. For FR2 there is not enough input to encourage/preclude anything.

The following was agreed:

* For FR1 companies are encouraged to provide simulation results for one-shot transmission

**Summary how to calculate the latency**

Based on companies input, the following two observations are made:

* Components from table 5.7.1.1.1.-1 for DL and table 5.7.1.1.2.-1 for UL from TR 37.910 are ok to be used as a starting point.
* In case re-tx is simulated, the alignment delay for the re-TX at the gNB side (which is not included in the tables from the TR 37.910) should also be added to the latency

Then, there is consensus on most values from TR 37.910 except for the UE and gNB processing times. Common values can then be agreed for all components except the processing times.

For the processing times, both the numbers from 37.910 and from TR 38.824 have been mentioned, with the numbers in TR 37.910 for a single link whereas the numbers in TR 38.824 should also somewhat take into account that multiple UEs have to be processed. Therefore, the numbers in 38.824 are more suitable for our case than the numbers in 37.910. However, even if the numbers in 38.824 are more realistic for a real system than the numbers in 37.910, it was understood during the URLLC SI, that also those values were only for obtaining aligned values for comparison across companies.

For the real system performance, it is very implementation dependent and these number should be reported by companies during our simulation campaign for 5G-ACIA.

The following was proposed:

Proposal 2: For the E2E latency, following assumptions are made:

* **Components from table 5.7.1.1.1.-1 for DL and table 5.7.1.1.2.-1 for UL from TR 37.910 are used to calculate the E2E latency** 
  + **In case re-tx is simulated, the alignment delay for the re-TX at the gNB side (which is not included in the tables from the TR 37.910) should also be added to the latency**
* **Companies report the UE processing delay and gNB processing delay, for other components, the values from table 5.7.1.1.1.-1 for DL and table 5.7.1.1.2.-1 for UL from TR 37.910 are assumed**

The proposal was supported by Intel, Huawei/HiSilicon, vivo, ITRI and ZTE. Ericsson preference was not to fork the assumptions and that all companies use the same assumptions (i.e., 37.910). Due to lack of time, no further discussions could not be done and the proposal is noted.

### 3.2.6 Additional clarifications of simulation assumptions

The following clarifications the simulation assumptions were proposed by companies:

* CSA metric with no consecutive errors is mandatory in round 2
* Narrow down to InF-DH explicitly
* Confirm options of traffic arrival for DL and UL
* Agree on number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled

Companies can give input on these proposals below. For the last issue, please provide input on what values should be used.

|  |  |
| --- | --- |
| Company | Input |
| vivo | What is the intention of addintionally introducing CSA metric with no consecutive errors? What is the defination of CSA metric with no consecutive errors?  In our understanding, CSA metric with no consecutive errors is same as PER. Is that common understanding?  If so, suggest to make the metric more general.  There is no need to agree on number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled, since they are heavily depend on simulation capability. Suggest leave it to company report. |
| HW/HiSi | *CSA metric with no consecutive errors is mandatory in round 2*  [HW/HiSi]: This is unclear to us. The definition of the CSA is in the LS from 5G-ACIA and has considered the survival time. A detailed formula is given there. Also, the distribution of CSA is one of the two agreed metrics for performance evaluation. What is the meaning of “CSA metric with no consecutive errors is mandatory”? Could you please clarify?  *Narrow down to InF-DH explicitly*  [HW/HiSi]:Ok  *Confirm options of traffic arrival for DL and UL*  [HW/HiSi]: For DL, we support 5G-ACIA options 1 and 3. For UL, both are reasonable, and Opt1 is slightly preferred.  *Agree on number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled*  [HW/HiSi]: We noticed that many simulations seem to be too short. In order to get reliable results, we would propose that about 100/(target error rate) packets should be generated. |
| ZTE | Our understanding on ‘CSA metric with no consecutive errors’ is to use the CSA metric defined in 5G ACIA LS. Anyway, it’s better to clarify the intention here.  We are fine to narrow down to InF-DH channel model only.  Regarding traffic arrival for DL and UL, we prefer 5G-ACIA Option 1 as agreed in round 1.  Limiting a certain number of samples could possibly resulting in less input from companies, though some companies may want to share their input for some metrics without relying on a large number of samples. So, we share with vivo that this could leave to companies report, while companies should explicitly report the number of samples in round 2. |
| Intel | Among the additional discussion points, it seems the number of samples may be important to agree. However, 100/(target error rate) may be nice but infeasible from computation time perspective for some sources. Consider 10/target error rate as a minimum set.  Note, that CSA is not equivalent to target PER, thus 10/(1-target CSA) may be more precise. |
| ITRI | *CSA metric with no consecutive errors is mandatory in round 2*  [ITRI]: The definition of the CSA is clearly defined in the LS from 5G-ACIA. “CSA metric with no consecutive” is similar to the PER, and we support the CSA metric with no-zero survival time (1ms in UC#2) as defined in the LS from 5G-ACIA, to satisfy the demand from 5G-ACIA.  *Narrow down to InF-DH explicitly*  [ITRI]:Ok  *Confirm options of traffic arrival for DL and UL*  [ITRI]: We support 5G-ACIA option 1 for DL and option 1 for UL as mandatory.  *Agree on number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled*  [ITRI] Suggest leave it to company report. |
| Nokia | 1. CSA metric with no consecutive errors is mandatory in round 2 🡺 OK to Nokia 2. Narrow down to InF-DH explicitly 🡺 OK to Nokia 3. Confirm options of traffic arrival for DL and UL 🡺 OK to Nokia to confirm October outcome, if just one is to be picked, would prefer option 1. 4. Agree on number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled 🡺 unnecessarily detailed level of agreement, but a recommendation could of course be discussed |
| Ericsson | CSA metric with no consecutive errors is mandatory in round 2  [Ericsson] We do not see the point of additionally reporting CSA metric with no consecutive errors. By definition, two or more consecutive packet errors are counted as system unavailable. This is already reflected in the existing CSA metric.  Narrow down to InF-DH explicitly  [Ericsson] Agree  Confirm options of traffic arrival for DL and UL  [Ericsson] Agree. That is, Option-1 for DL traffic and Option-1 for UL traffic relationship to DL.  Agree on number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled  [Ericsson] OK to leave to companies’ choice. On the other hand, these should be reported by each company when submitting simulation results. |

**Conclusion on these issues:**

* No consensus on CSA metric with no consecutive errors is mandatory
* Narrow down channel model to InF-DH explicitly
* Option-1 for DL traffic and Option-1 for UL traffic relationship to DL is still baseline. Additional results can be submitted
* Number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled are left to companies’ choice

## 3.7 Format for inputs to week

In order to generate the final report for the study, it would be good to align the input formats. The following is recommended provided:

* CDF of packet error rate for UL and DL
* CDF of CSA for UL and DL
* Tabulated values for percentage of UEs satisfying 1ms latency and 99.9999% reliability/CSA requirement for each simulated case
* CDF for coupling loss and geometry for calibration

# 4 Conclusions

This is a summary of the discussion in week 2 of the 5G-ACIA discussion. The input contributions containing the simulation results can be found in the reference section. The conclusions and agreements from the discussions are listed below.

Conclusion on colleting simulation assumptions:

* The final Excel sheet can be found [here](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Final%20Summary/Simulation%20assumptions%20for%20calibration%20Final.xlsx).

Conclusion on FR2 antenna assumptions:

* 2RX/TX is still the baseline
* Results for additional configurations can be provided

Conclusion on cell coordination:

* No coordination is baseline
* Results with cell coordination can be provided

Conclusion on MU-MIMO:

* SU-MIMO is baseline
* Results with MU-MIMO can be provided

Agreement for latency:

* For FR1 companies are encouraged to provide simulation results for one-shot transmission

Noted proposal for latency

* For the E2E latency, following assumptions are made:
  + Components from table 5.7.1.1.1.-1 for DL and table 5.7.1.1.2.-1 for UL from TR 37.910 are used to calculate the E2E latency
    - In case re-tx is simulated, the alignment delay for the re-TX at the gNB side (which is not included in the tables from the TR 37.910) should also be added to the latency
  + Companies report the UE processing delay and gNB processing delay, for other components, the values from table 5.7.1.1.1.-1 for DL and table 5.7.1.1.2.-1 for UL from TR 37.910 are assumed
* Supported by Intel, Huawei/HiSilicon, vivo, ITRI, ZTE

Conclusion on additional simulation assumptions:

* No consensus on CSA metric with no consecutive errors is mandatory
* Narrow down channel model to InF-DH explicitly
* Option-1 for DL traffic and Option-1 for UL traffic relationship to DL is still baseline. Additional results can be submitted
* Number of samples, minimum number of packets per UE and minimum number of UEs / network drops modelled are left to companies’ choice

Conclusions on format for submissions to round 2

* Companies will provide
  + CDF of packet error rate for UL and DL
  + CDF of CSA for UL and DL
  + Tabulated values for percentage of UEs satisfying 1ms latency and 99.9999% reliability/CSA requirement for each simulated case
  + CDF for coupling loss and geometry for calibration

# References

1. [RP-202069](https://protect2.fireeye.com/v1/url?k=41a5db26-1f051960-41a59bbd-86fc6812c361-73f443258ff773bf&q=1&e=bc078f84-983d-45f3-ab31-19e60d911036&u=https%3A%2F%2Fwww.3gpp.org%2Fftp%2Ftsg_ran%2FTSG_RAN%2FTSGR_89e%2FDocs%2FRP-202069.zip), “Way forward on RAN work for 5G ACIA requested simulations“, Ericsson
2. “[Simulation results for 5G-ACIA in the first round](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/HwHiSi%20-%20Simulation%20results%20for%205G-ACIA%20in%20the%20first%20round.docx) Huawei, HiSilicon
3. “[5G-ACIA LS – Phase 2 input](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/INTEL%20-%205G-ACIA%20LS%20-%20Phase%202%20inputs.docx)”, Intel Corporation
4. “[Simulation Assumptions and URLLC Performance Evaluations for 5G-ACIA Performance Evaluation Round 1](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/ITRI_5G%20ACIA%20Simulation%20Result%20for%20InF-DH%204GHz.docx)”, ITRI
5. “[First round of simulation results for 5G-ACIA evaluation](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/NOKIA%20-%205G-ACIA%20First%20round%20of%20simulation%20results.zip)”, Nokia, Nokia Shanghai Bell
6. “[First round of FR1 simulation results for 5G ACIA URLLC LS response](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/QUALCOMM-5G-ACIA_URLLC_simulation_results_1st_round_FR1.docx) ”, Qualcomm CDMA Technologies
7. “[Simulation Assumptions and URLLC Performance Evaluations for 5G-ACIA Performance Evaluation Round 1](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/Qualcomm5G-ACIA_URLLCResultsRound1_FR2.docx)”, Qualcomm CDMA Technologies
8. “[5G-ACIA 1st round URLLC evaluation results](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/vivo-5G-ACIA%201st%20round%20URLLC%20evaluation%20results.DOCX)”, vivo
9. “[Simulation Results for 5G-ACIA (First round)](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/Ericsson%205G-ACIA%20Simulation%20Results%20Round1.zip)”, Ericsson
10. “[ZTE-5G-ACIA evaluations - 1st round of simulation results](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/ZTE-5G-ACIA%20evaluations%20-%201st%20round%20of%20simulation%20results.docx)”, ZTE
11. “[INTEL - 5G-ACIA LS - Phase 2 inputs - updated](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Company%20Inputs/INTEL%20-%205G-ACIA%20LS%20-%20Phase%202%20inputs%20-%20updated.docx)”, Intel
12. “[Excel Sheet for simulation assumptions calibration](https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Inbox/Drafts/5G-ACIA%20December/Final%20Summary/Simulation%20assumptions%20for%20calibration%20Final.xlsx)”