**3GPP TSG RAN WG1 #104-e R1-20xxxxx**

**e-Meeting, Jan 25th – Feb 5th, 2021**

**Agenda item:** 8.14.1

**Source:** Moderator (Qualcomm)

**Title:** Email discussion/approval for traffic model and capacity KPI for XR

**Document for:** Discussion and Decision

# Introduction

[104-e-NR-XR-01] Email discussion/approval for traffic model and capacity KPI – Eddy (Qualcomm)

Below are the agreements made during the RAN1 meeting on Thursday, Jan 28th 2021. In this document, we continue to discuss open topics on traffic model and capacity KPI for XR.

|  |
| --- |
| Agreements**:** RAN1 adopts a parameterized statistical traffic model for evaluation of XR and CG, and KPI with details as shown below (RAN1 strives to agree on the remaining details during RAN1 #104e, based on SA4 input):   * There are M1 and M2 streams in DL and UL respectively   + At least adopt the case where M1=1 & M2=1   + FFS the values of M1 and M2, including the possibility of being application-dependent * DL   + Bitrate for video streaming     - VR/AR: [60 Mbps (mandatory), 30 Mbps (optional)]     - CG: [30 Mbps (mandatory), 45 Mbps (optional)]       * FFS: other optional values   + Air interface Packet Delay budget (PDB)     - Air interface delay is measured from the point when a packet arrives at gNB to the point when it is successfully delivered to UE     - Air interface PDB for video streaming       * VR/AR: [10ms (mandatory), 20ms (optional)]       * CG: [15ms (mandatory), 30ms (optional)]         + FFS: other optional values   + FFS: Frame-level/IP packet-level modeling for packet arrival, latency measure, etc.   + FFS: Packet size, including the possibility of varying packet sizes   + FFS: Packet Inter arrival time including the possibility of modeling jitter * UL   + FFS: Bitrate   + FFS: Air interface Packet Delay budget (PDB)   + FFS: Frame-level/IP packet-level modeling for packet arrival, latency measure, etc.   + FFS: Packet size * Per UE KPI   + Baseline: A UE is declared a satisfied UE if more than X (%) of packets are successfully transmitted within a given air interface PDB. The exact value of X is FFS.   + FFS: In addition to the baseline, the following additional method is FFS     - When determining a XR/CG user is satisfied or not, the following factors are considered. FFS how to use those factors.       * Packet loss information       * Packet delay information       * Some XR/CG source related information if they can be available within RAN, e.g. the mapping between packet and slices or frames and the packet importance       * Multiple data streams traffic model   + FFS if there are multiple streams (if adopted) * FFS additional aspects not addressed above. * Note 1: Companies are encouraged to provide details such as parameters (e.g., mean, STD, etc.), distributions, etc., by analyzing SA4 input, e.g., V/S/P traces * Note 2: All FFS points above are to be further discussed in RAN1 #104e |

# 2nd Round Email Discussion

DL Video Stream

### Single Flow DL Video Stream

**In this section, our discussion is limited to the case of single flow DL video streaming**. Multi-flow/stream aspects such as I-frame, P-frame, audio, data stream are discussed in next sections.

**Statistical traffic model for DL video stream**

From GTW discussion, it seems that the statistical traffic model for DL video stream in Figure 1 is aggregable, where a packet is assumed to represent multiple IP packets corresponding to a single video frame for modelling/evaluation purposes, e.g., traffic arrival, packet size, evaluation of latency and reliability. Given the diverse views on terminology of “frame level” vs. “V-trace level” vs. “P-trace level”, etc., these terms will not be used in the subsequent discussion.

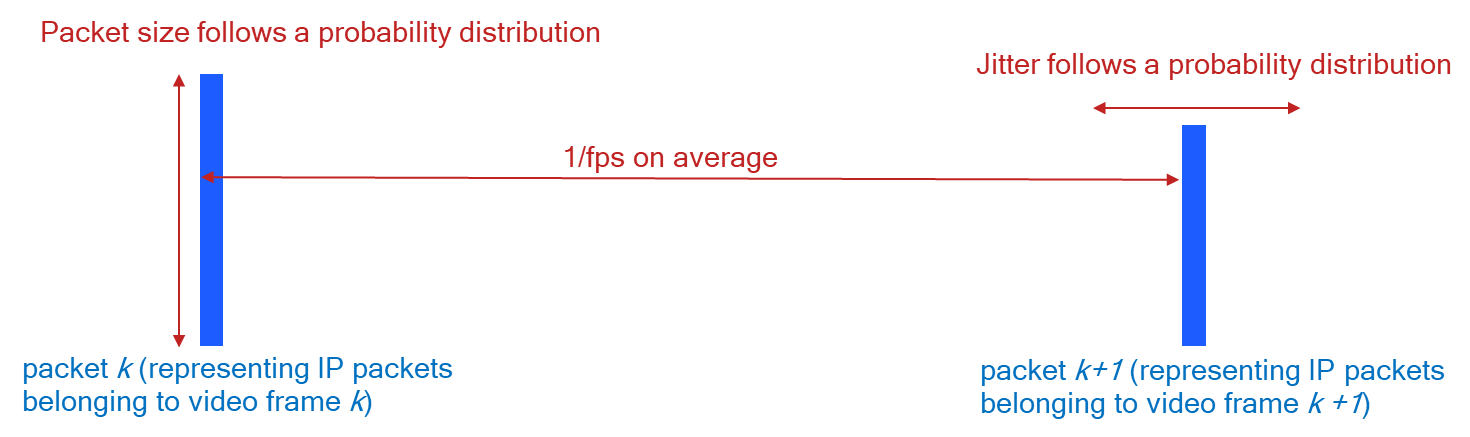


Figure 1. Statical traffic model for DL video stream

**Proposal 1**. The statistical traffic model for a single DL video stream in Figure 1 is adopted.

**Question 1**. Can we agree on Proposal 1? Please share your view.

|  |  |
| --- | --- |
| Company | View |
| OPPO | We are ok with proposal 1 |
| Ericsson | Support |
| ZTE,Sanechips | Agreed |
| Xiaomi | Agreed |
| CATT | We would like to clarify the assumption of packet generation in the proposed video traffic model. We are OK in modelling the packet inter-arrival based on statistic model. However, we need to define the traffic model with clarification on the assumption   * Whether this is a single user traffic model or multi-user traffic models at the gNB buffer   + If this is single user traffic model, the packet arrival at the gNB buffer should follow the packet generation based on video stream source codec defined by SA4 or H.264/265/266. Whether the packet size is based on actual P-trace or V-trace model or statistical model.   + If multi-user traffic model is used, the packet size is model as the statistic distribution of aggregated traffic arrival from multiple users. The statistic distribution is based on the sum of multiple random variables with each random variable is the video stream generated from traffic model by SA4 or H.264/265/266. * We need to determine whether each video stream has same QoS for different frame types in a video stream. If more than one QoS level are defined, there would be more than one priority queue at the gNB scheduler, we need to model the packet arrival at each priority queue. |
| vivo | Fine with proposal 1. |
| Nokia, NSB | Agree with Proposal 1. |
| Huawei, HiSilicon | Support. We think the following red sentence (copied from FL’s explanations above) seems good for clarification, and suggest to be added.  **Proposal 1**. The statistical traffic model for a single DL video stream in Figure 1 is adopted, where a packet is assumed to represent multiple IP packets corresponding to a single video frame for modelling/evaluation purposes, e.g., traffic arrival, packet size, evaluation of latency and reliability. |
| Futurewei | Support Proposal 1. |
| Sony | Support proposal 1 |
| QC | We support Proposal 1. |
| LG | Yes, we can agree on Proposal 1. |
| InterDigital | Ok with proposal |
| Apple | In our view, this is about the video stream model, it is not coupled with single flow or multiple flow discussion. For AR2, for example, merging audio/data stream to the video stream can lead to latency bound cannot be met. |
| MTK | Solved in GTW |

**Frame per second (fps)**

Summary of 1st round email discussion (for DL video stream)

* 60 fps **(9)**: MTK, QC, Apple, Futurewei, Interdigital, DCM, ZTE, vivo, HW
* ~~120 fps~~ **~~(1)~~** ~~: OPPO~~
* Both 60 fps and 120 fps **(2)**: Xiaomi, AT&T, OPPO
* Both 30 fps and 60 fps **(1)**: Sony
* 120 fps optional: DCM, HW
* 90 fps optional: HW
* FTP Model 3: Nokia, CATT, LG
  + Note: This proposal was made before the statistical model in Figure was accepted by these companies. So, I assume now that these companies can accept defining fps values.

For video stream, most of companies are proposing 60 fps, while some companies address other values such as 30 fps, 90 fps, and 120 fps. Based on this, the following proposal is made.

**Proposal 2**. For a single DL video stream for AR/VR/CG, 60 fps is the baseline. Other values, e.g., 30, 90, 120 fps can be optionally evaluated.

**Question 2**. Please share your view on Proposal 2.

|  |  |
| --- | --- |
| Company | View |
| OPPO | Our position is not captured correct in the above summary. I make a correction that we support both 60 fps and 120 fps.  We prefer to keep both 60 and 120 fps as the baseline. We also can accept 120fps as optional if the majority view support only 60 fps as the baseline |
| Ericsson | Support. |
| ZTE,Sanechips | Agreed |
| Xiaomi | We are fine with FL proposal. |
| CATT | We are OK that the average inter-arrival time is 16.67 ms (60 fps) for each video stream. It will depends whether single user or multi-user traffic arrival at gNB buffer is modelled. |
| Vivo | Fine with proposal 2. |
| Nokia, NSB | Agree with Proposal 2. |
| Huawei, HiSilicon | Support |
| Futurewei | Support Proposal 2. |
| QC | Support |
| LG | We are okay with Proposal 2. |
| InterDigital | Ok with proposal |
| Apple | It is quite clear from SA4 input, there are multiple flows for AR2. The video stream modelling aspects should be decoupled with single flow vs multiple flow discussion. The proposal can be modified as:  For ~~a single~~ DL video stream for AR/VR/CG, 60 fps is the baseline. Other values, e.g., 30, 90, 120 fps can be optionally evaluated. |

**Data rate for DL video stream**

Summary of 1st round email discussion

* **(7)** MTK, Futurewei, DCM, vivo, QC, LG, Xiaomi
  + VR/AR: 60Mbps
  + CG: 30 Mbps
* **(3)** Nokia, InterDigital, AT&T
  + VR/AR: 60Mbps
  + CG: 30, 45 Mbps
* **(2)** CATT, Sony
  + AR/VR: [30, 60] Mbps
  + CG: [8, 30] Mbps
* **(1)** ZTE
  + VR/AR: 45Mbps
  + CG: 45 Mbps
* **(1)** Ericsson
  + AR/VR: 30 Mbps (mandatory), 60 Mbps (optional)
  + CG: 8 Mbps (mandatory), 30 Mbps (optional)
* **(1)** OPPO
  + VR/AR: 90 Mbps, 45 Mbps
  + CG: 12 Mbps, 45 Mbps

Based on 1st round email discussion, the following proposal is made.

**Proposal 3**.

* Data rate for DL video stream:
  + VR/AR: 60 Mbps
  + CG: 30 Mbps

Other values can be optionally evaluated.

Note: It is clarified that the data rate indicates the arrival bitrate at gNB MAC layer.

**Question 3**. Please share your view on Proposal 3.

|  |  |
| --- | --- |
| Company | View |
| OPPO | I am not against the media bitrate of 60/30. I am just wondering why we don’t use the model of SA4 that the bitrate from cloud/edge to gNB is 1.5 times of media bitrate. Could some companies or FL clarify it? |
| Ericsson | Don’t support. In the SA4 LS, the following aggregated data rates are mentioned:   * CG: 8Mbps or 30Mbps * AR2/VR2: 30Mbps   We believe that it is important to also investigate data rates lower than 30Mbps – lower bitrates would be the result of rate adaptation in the device.  Furthermore, lower bitrates may trigger additional bottlenecks, thereby providing input to future improvements.  If/when SA4 provides additional input, resulting in higher bitrates, RAN1 should discuss to also include those.  We are also trying to understand the SA4 input on the “1.5”. |
| ZTE,Sanechips | If we look at OPPO and Ericsson’s input together, we would see the challenging situation for evaluation is 30\*1.5 = 45Mbps, which is our proposal. |
| Xiaomi | We are fine with FL proposal. A clarification: the data rate in the note is the average data rate, right? |
| C**ATT** | We are OK with FL proposal. However, we would like to know whether the video stream has only one QoS level or multiple QoS levels. |
| Vivo | Fine with proposal 3.  On Xiaomi’s comment, we think it is the average data rate. |
| Nokia, NSB | We generally support Proposal 3. We also propose to add CG optional value 45 Mbps (referring to 4K video) so that companies choosing the optional values can actually compare the results.   * Data rate for DL video stream:   + VR/AR: 60 Mbps (aggregated video traffic to the XR device)   + CG: 30 Mbps     - Optional: 45 Mbps   Other values can be optionally evaluated.  Note: It is clarified that the data rate indicates the arrival bitrate at gNB MAC layer.  We would also like to clarify that for VR/AR 60 Mbps refers to the aggregated traffic to the XR device |
| Huawei, HiSilicon | Since other values can be optionally evaluated, we are generally fine with the values in the proposal. We think the current values refer to 60 fps, right? So adding “@60fps” as below may be more accurate since the data rate at 120 fps could be different.  We suggest to remove the last Note. The data rate here is used to derive the packet size distribution, i.e., as shown in Proposal 5, Mean = (data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]. And RAN1 can further do evaluation based on such packet size distribution. So the Note here does not bring any new information, but may cause confusion that companies may guess what else information could the Note tell.  **Proposal 3**.   * Data rate for DL video stream:   + VR/AR: 60 Mbps @60 fps   + CG: 30 Mbps @60 fps   Other values can be optionally evaluated.  ~~Note: It is clarified that the data rate indicates the arrival bitrate at gNB MAC layer.~~ |
| Futurewei | We are ok with Proposal 3. |
| QC | We support. It would be clearer if it is clarified that given rates are **average** bit rate. |
| LG | We are okay with Proposal 3. |
| InterDigital | Ok with proposal, assuming these are average data rates |
| Apple | We don’t support this proposal; we need to follow SA4 inputs. |
| MTK | Solved in GTW |

**Packet size distribution**

Summary of 1st round email discussion on distribution of packet size

* **(13)** Truncated Gaussian: OPPO, MTK, Xiaomi, QC, CATT, Futurewei, InterDigital, DCM, AT&T, vivo, HW, Sony, Ericsson
* **(2)** Constant IP packet size: Nokia, LG
* **(1)** ZTE: need to first decide on Gaussian vs. Pareto

In addition, Figure 2 compares Truncated Gaussian distribution vs. distribution of “sum of IP packet sizes belonging to one video frame in SA4 P trace”. The left hand figure is for CBR and the right hand figure is for VBR. These results confirm that Truncated Gaussian distribution is a good approximation for packet size distribution according to the statistical model in Figure 1.

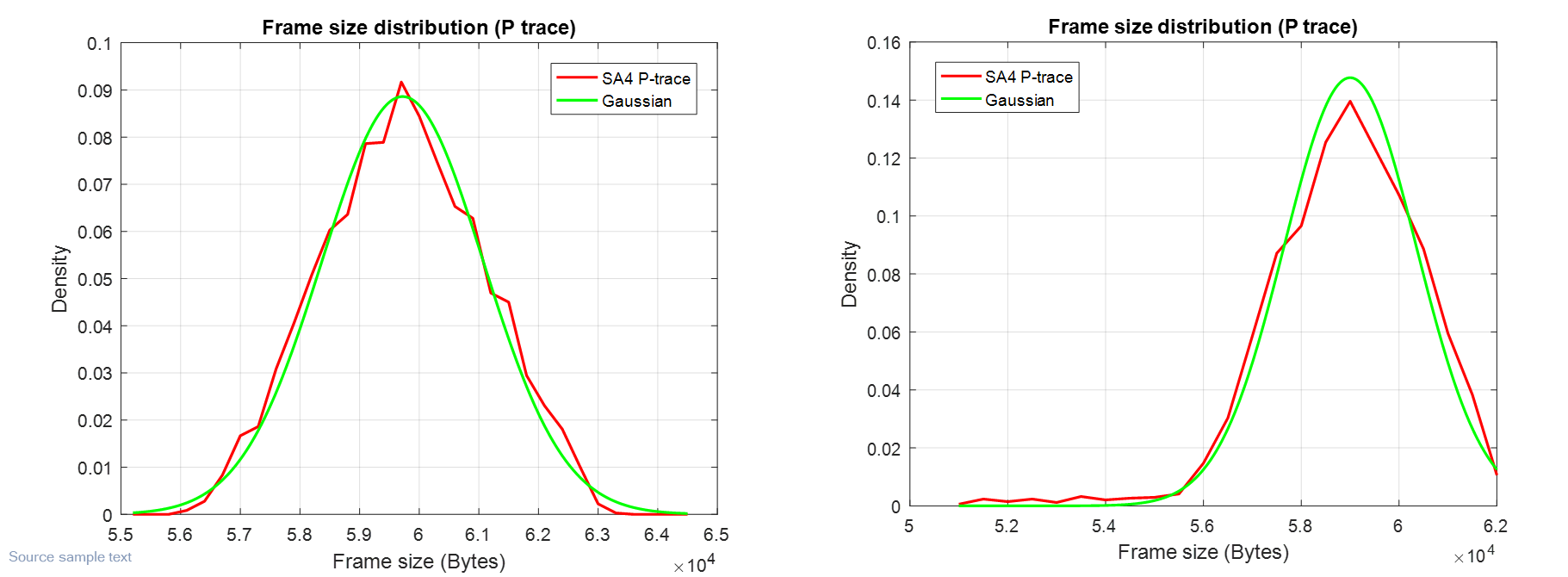


Figure 2. Distribution of sum of IP packet sizes belonging to one video frame per SA4 P-trace vs. Gaussian distribution with the same STD

Based on companies’ view in 1st round email discussion and based on the statistics from SA4 P-trace, the following proposal is made.

**Proposal 4**. Truncated Gaussian distribution is used for the packet size distribution of a single video stream for AR/VR/CG.

**Question 4**. Please share your view on Proposal 4.

|  |  |
| --- | --- |
| Company | View |
| OPPO | Support |
| Ericsson | Support |
| ZTE,Sanechips | OK for progress. We have to point out in the SID, it’s explicitly captured that Pareto distribution as starting point. This can be reflected with the addition in red.  **Proposal 4**. Truncated Gaussian distribution is used for the packet size distribution of a single video stream for AR/VR/CG.  Truncated Pareto distribution could also be considered. |
| Xiaomi | Fine with FL proposal. |
| CATT | We are OK with the proposal. However, we like to clarify whether this is single user or multi-user traffic model. |
| Vivo | Fine with proposal 4. |
| Nokia, NSB | Agree with proposal 4. |
| Huawei, HiSilicon | Support |
| Futurewei | Support Proposal 4. |
| QC | Support |
| LG | As we agreed on Proposal 1, our preference on the constant IP packet size doesn’t hold any more. We are okay with Proposal 4 in favour of the progress. |
| InterDigital | Support |
| MTK | Solved in GTW |

**Parameters of Truncated Gaussian distribution for Packet size**

Summary of 1st round email discussion: Parameters of truncated Gaussian distribution for packet size,

* Mean: (data rate) / (# files per second, i.e., fps) / 8 [bytes]
* STD
  + QCOM: STD: 13% of Mean
  + MTK: 13% of Mean
  + Vivo: 12.5% of Mean
  + Ericsson: 30% of Mean
* Max packet size
  + QCOM 1.44 x Mean
  + MTK: 1.44 x Mean
  + Vivo: 1.563 x Mean
  + Ericsson: 2 x Mean
* Min packet size
  + Vivo: 67 bytes
  + Ericsson: 0 bytes

**Proposal 5**. If truncated Gaussian distribution is used for the packet size distribution of a single video stream for AR/VR/CG, based on the 1st round of email discussion summarized above, the following parameter values are adopted.

* Mean: Derived from data rate and fps as follows.
  + (data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
  + Note: Data rate is the arrival bitrate at gNB MAC layer.
* STD
  + 15% of Mean
* Max packet size
  + 1.5 x Mean
* Min packet size
  + 67 bytes

**Question 5**. Please share your view on Proposal 5.

|  |  |
| --- | --- |
| Company | View |
| Ericsson | We are essentially ok, but if the distribution is not symmetric around the mean, the resulting data rate will be different (lower) from what is assumed. |
| ZTE, Sanechips | What matters to a truncated Gaussian model is the mean/std and max packet size. No need to agree to the min packet size which has marginal impact to the model. Alternatively we could use something like min = mean – 3\*std. |
| CATT | We are OK on the proposal as working assumption. We need to further study how this is associated with SA4 XR traffic model |
| vivo | Fine with proposal 5. |
| Nokia, NSB | **Mean** – support  **STD** – please clarify the choice of 15%, not clear  **Max and min:**  We do not see the values presented in Proposal 5 represent the Figure 2 data and our previous analysis of single packets in CG. We think the min packet burst size from Proposal 5 is too small. Given the average **single** packet size (CG analysis) is 1200 bytes, 67 bytes for a minimum **burst** of packets **sounds unrealistic**. Maximum burst of packets equal to 1.5xMean seems much larger than what we observe from Figure 2.  We propose to continue the iteration on the values to be more reasonable in our choice. In case we will not be able to find reasonable values during this meeting (since not much time is left) we propose to put the values in square brackets for simulation purposes and add **FFS: clarify the values during 105-e**. |
| Huawei, HiSilicon | Based on our analysis of the SA4 outcome, I frame size could be 2~3 times of P frame. So we suggest to change the max packet size as “~~1.5~~3 x Mean” to reflect this.  As for the min packet size, it seems “67 bytes” is obtained from some specific video sources and may be not so general. So we suggest to take “0.1 x mean” to be more general.  Similar to our reply for Proposal 2, we suggest to remove the Note.  **Proposal 5**. If truncated Gaussian distribution is used for the packet size distribution of a single video stream for AR/VR/CG, based on the 1st round of email discussion summarized above, the following parameter values are adopted.   * Mean: Derived from data rate and fps as follows.   + (data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]   + ~~Note: Data rate is the arrival bitrate at gNB MAC layer.~~ * STD   + 15% of Mean * Max packet size   + ~~1.5~~3 x Mean * Min packet size   + ~~67 bytes~~0.1 x Mean |
| Futurewei | We are ok with Huawei/Hisi’s revision. |
| QC | We are fine with the proposal. From our point of view, the first step would be to agree on the type of distribution we use. In case companies cannot agree on the specific parameter values of the model, we propose to agree them with square bracket for progress purpose. |
| LG | For those values in Proposal 5, brief summary of how the FL picked up those values would be helpful. |
| InterDigital | Generally ok with the model described in the proposal. However, clarification would be needed for the proposed std, min/max values in terms of how these relate to the values provided by SA4 |
| Apple | Suggested wording change:  If truncated Gaussian distribution is used for the packet size distribution of ~~a single~~ DL video stream for AR/VR/CG, based on the 1st round of email discussion summarized above, the following parameter values are adopted. |
| MTK | Fine with FL proposal, while we do not think “Min packet size” is needed. |

**Jitter modelling in packet arrival time**

**Proposal 6**. If the statistical model in Figure 1 is agreed for a single DL video stream, a truncated Gaussian distribution is used for jitter modelling, i.e., for X fps (e.g., X = 60), arrival time of packet k = k / X + *J*, where is *J is* a random variable drawn from a truncated Gaussian distribution.

**Question 6**. Please share your view on Proposal 6.

|  |  |
| --- | --- |
| Company | View |
| OPPO | Support |
| Ericsson | Support |
| ZTE,Sanechips | **Proposal 6**. If the statistical model in Figure 1 is agreed for a single DL video stream, a truncated Gaussian distribution is used for jitter modelling, i.e., for X fps (e.g., X = 60), arrival time of packet k is k / X + *J*, where is *J is* a random variable drawn from a truncated Gaussian distribution.  The equal symbol is quite confusing, which could be addressed by the change in red. It should be clarified that J should be measured in s, not in ms as implied by P7. |
| Xiaomi | Support. A clarification: this does not preclude the possibility to add additional offset to randomize the arrival time among different UEs, right? |
| CATT | We are OK modelling packet inter-arrival time at gNB using truncated Gaussian model. However, the inter-arrival time is a stochastic process and depends on whether traffic is modelled as single user or multi-user traffic arrival at the gNB buffer.   1. For single user single session traffic model with one hop delay jitter – the inter-arrival time is exponential distribution 2. For single user multi-session traffic model with one hop delay jitter – the inter-arrival time is Poisson distribution 3. For single user traffic model with multi-hop delay jitter – the inter-arrival is truncated Gaussian distribution. |
| Vivo | Fine with proposal 6. |
| Nokia, NSB | Support |
| Huawei, HiSilicon | Share similar view with ZTE. We suggest the following red changes.  **Proposal 6**. If the statistical model in Figure 1 is agreed for a single DL video stream, a truncated Gaussian distribution is used for jitter modelling, i.e., for X fps (e.g., X = 60), arrival time of packet k is k / X\*1000 [ms] + *J* [ms], where is *J is* a random variable drawn from a truncated Gaussian distribution. |
| Futurewei | Support Proposal 6. |
| QC | We support this model (with clarification on time unit of J).  The given equation should be the case when there is 0 offset for the traffic arrival. In case there is multiple Ues in the cell, each UE could have different traffic arrival offset which is later added to each UE’s arrival time. |
| LG | Okay with Proposal 6 in principle. Some clarifications on the unit of the arrival time and the index k would be helpful. |
| InterDigital | Ok with proposal |
| Apple | **Suggested wording change:**  **Proposal 6**. If the statistical model in Figure 1 is agreed for ~~a single DL~~ DL video stream, a truncated Gaussian distribution is used for jitter modelling, i.e., for X fps (e.g., X = 60), arrival time of packet k = k / X + *J*, where is *J is* a random variable drawn from a truncated Gaussian distribution. |
| MTK | Support the proposal |

**Parameters of Truncated Gaussian Distribution for jitter in packet arrival time**

Summary of 1st round email discussion:

* Mean
  + Zero: MTK, QC
* STD
  + MTK: 3 ms
  + QC: sqrt(2)
* Range
  + MTK: [-7, 7] ms
  + QC: [-4, 4] ms
  + Ericsson: max < 7ms

Based on 1st round email discussion, the following proposal is made.

**Proposal 7**. If truncated Gaussian distribution is agreed for jitter modelling, the following parameters are adopted.

* Mean
  + 0
* STD
  + 2 ms
* Range
  + [-4, 4] ms

**Question 7.** Please share your view on Proposal 7.

|  |  |
| --- | --- |
| Company | View |
| OPPO | In the LS from SA4, it is said that “Pre-encoding delay: Encoder pre-delay is varying between 10 to 20ms”. According SA4’s model, the range [-4, 4] is not sufficient to reflect the varying delay. Compared to [-4, 4], the range [-7, 7] seems better. |
| Ericsson | We can support, although conceptually a negative jitter seems strange. |
| CATT | The values are too small and not realistic. We need further study on the network delay jitter with more realistic model. |
| vivo | Fine with the proposal. |
| Nokia, NSB | It would be good to clarify the technical reasoning/references for numerical values (e.g., 2 ms STD, [-4,4] ms for range). |
| Futurewei | We are ok with the proposal. |
| QC | We think this could be a good starting point. We want to note that the truncation bound of jitter should be carefully chosen so that it does not introduce out of packet arrival. |
| LG | If the encoder pre-delay varies b/w 10 to 20ms, then we can also consider [-5, 5] as it corresponds to the SA4 input. |
| InterDigital | We are ok with the values as a starting point |
| MTK | Our proposed value: STD 3ms ms is based on Google stadia measurement, but we are fine with FL proposed values and [-5,5]/[-4,4]/[-7/7] range as starting point. |

**Air interface Packet Delay Budget (PDB)**

The follow agreements were made in the last GTW session.

|  |
| --- |
| * Air interface Packet Delay budget (PDB)   + Air interface delay is measured from the point when a packet arrives at gNB to the point when it is successfully delivered to UE   + Air interface PDB for video streaming     - VR/AR: [10ms (mandatory), 20ms (optional)]     - CG: [15ms (mandatory), 30ms (optional)]       * FFS: other optional values |

Summary of 1st round email discussion on Air interface PDB

* VR/AR:
  + 5ms (1) : HW
  + 10ms (13) : MTK, vivo (VR), QC, Xiaomi, Nokia, CATT, InterDigital, DCM, Nokia, HW, sony, LG, Ericsson
  + 15ms (1): Futurewei
  + 20ms (6): Nokia, CATT, DCM (optional), Nokia, vivo(AR), sony
  + 30ms: HW
* CG
  + 10ms (1): Xiaomi,
  + 15ms (13): MTK, vivo, QC, Nokia, Futurewei, CATT, InterDigital, DCM, vivo, HW, sony, LG, Ericsson
  + 30ms (5): Nokia (optional), CATT, CDM (optional), Nokia (optional), HW, sony
  + 50ms (1): HW
* ZTE: min (Delay for RAN1, remaining PDB for RAN1)

Based on 1st round email discussion, the following proposal is made.

**Proposal 8**. For a single flow DL video stream, the following air interface PDB values are evaluated, where PDB is the latency budget within which a packet has to be delivered to UE from the point when the packet arrive at gNB to meet the latency requirement.

* Air interface PDB for DL video stream
  + VR/AR: 10ms
  + CG: 15ms

Other values can be optionally evaluated.

**Question 8.** Please share your view on Proposal 8.

|  |  |
| --- | --- |
| Company | View |
| Ericsson | Support |
| Xiaomi | We still prefer to 10 ms PDB for all applications, but we can accept FL proposal. |
| CATT | We are OK with the proposal as the working assumption. We can re-visit the number next time when the detail of SA traffic models are considered. |
| Vivo | Support proposal 8.  If jitter is modelled, the impact of jitter on Air interface PDB need to be further clarified. |
| Nokia, NSB | Support |
| Huawei, HiSilicon | For video stream in a given XR or CG application, different roundtrip interaction delays can result in different user experiences. To reflect different levels of user experience, more values should be evaluated.  So the following red values are suggested.   * Air interface PDB for DL video stream   + VR/AR: 5ms, 10ms , 20ms   + CG: 10ms, 15ms, 30ms |
| Futurewei | Support Proposal 8. |
| QC | We support the PDB. We think the impact of jitter on packet expiration time needs to be clarified. What we want to capture is as follow.   * If the jitter of a packet is negative (i.e., packet arrives earlier than its expected time), then, the packet has extended time to be served.   If the jitter of a packet is positive (i.e., packet arrives late than its expected time), then the packet has shorter time to be served. |
| LG | Okay with Proposal 8. |
| InterDigital | Ok with the proposal |
| Apple | This is applicable to the video stream only. SA4 has provided latency requirement for other streams. |
| MTK | For VR/AR/CG, SA4 suggested two PDB values of 10/20ms. We share same view with Xiaomi to apply 10 ms PDB for all applications is better, but we can accept FL proposal. |

DL Multiple Streams/Flows

We have discussed statistical traffic model and detailed parameters for a single video stream in DL. Some companies are proposing to evaluate additional DL streams/flows such as I-frame, P-frame, audio, two eye buffers, etc. It does not seem to be very clear how to model and evaluate multiple streams/flows. Therefore, in this section, we are asking proponents of multiple DL streams/flow to present more details on top of the framework and detailed parameters for a single DL video stream discussed in the previous section.

**Question 9.** If you propose to evaluate multiple streams/flows such as I-frame, P-frame, audio, two eye buffers, etc. in DL, please present details of traffic model and how they can be evaluated referring to the framework and details parameters discussed for a single DL video stream.

|  |  |
| --- | --- |
| Company | View |
| vivo | Further study traffic model for multiple streams/flows in RAN1 #105. |
| Huawei, HiSilicon | Multiple data streams should be considered to model real XR applications more accurately. Each data stream may have different traffic characteristics (e.g., periodicity, data rate, packet size, etc.) and QoS requirements in DL/UL, e.g., FOV stream and non-FOV stream, video stream and audio stream, I-frame stream and P-frame stream, etc. If RAN1 only considers one data stream, it seems there is almost no difference between XR and URLLC traffic, and RAN1 is just going to evaluate URLLC traffic with more challenging requirements instead of XR traffic. Without multi-stream results, the SI and TR would not be assessing performance against the KPIs under assumptions which are a better match for actual XR applications.  As a starting point, I-frame stream and P-frame stream can be considered for the video stream. The details are given as follows:  **fps**  In video coding, a GOP (group of pictures) structure specifies the order in which intra- and inter-frames are arranged [1]. For example, assume GOP=60 and only consider I frame and P frame, then the frame pattern is IPP…PPI, i.e., the distance between two I frames is GOP=60 frames. So assume the single video stream’s fps is W, then the fps of I frame and P frame can be estimated as and, respectively.  [1] <https://en.wikipedia.org/wiki/Group_of_pictures>  **Packet size**  Based on our analysis of the SA4 outcome, I frame size could be 2~3 times of P frame. Let q denotes the ratio of I frame size and P frame size, i.e., , and assume the single video stream’s data rate is R. Then, the data rate of I frame and P frame can be estimated as and, respectively.  **PER, PDB**  It is common understanding that I frame is more important than P frame, since I frame is the reference frame for subsequent P frames. So the PER requirement of I frame should be more stringent than that of P frame. The PDB of I/P frame can the same as single video stream.  In summary, the following proposal is suggested:  **Proposal: The video stream in DL and UL is modelled as I-frame stream and P-frame stream, with the following details:**   |  |  |  |  | | --- | --- | --- | --- | |  | **I-frame stream** | **P-frame stream** | **Note** | | Data rate |  |  | * R: data rate of single video stream * GOP (Group of pictures): GOP=60   + Other values can be optionally evaluated. * q: ratio of I frame size and P frame size (), q=3   + Other values can be optionally evaluated. | | fps |  | W | * W: fps of single video stream | | Packet size distribution | Derived in the same way as single video stream | |  | | PDB | Same as single video stream | |  | | Packet success rate | 99.9% (mandatory), 99.5% (optional) | 99% (mandatory), 95% (optional) | Other values can be optionally evaluated. |  * **NOTE: each stream is based on the single video-stream model from Proposal 1** * **Note: other streams can be further discussed** |
| InterDigital | We think the evaluation of multiple streams for XR applications can be simplified by using certain QoS rules in addition to the QoS requirements (e.g. PDB, PER) of the individual streams belonging to an application. For example, if at least one of the streams (e.g. I-frame stream) in the group of streams is unable to meet its PDB requirement then the QoS requirement for the application can be considered to be not satisfied. |
| Apple | Downlink multiple flows should be modelled for AR2. It is clear SA4 includes multiple flows in the LS to RAN1. Moe specifically, audio stream, data stream and video stream are present in DL.  Assume video stream at 60 fps. The following timing diagram can be used to check the extra-latency experienced by the audio stream if audio packets wait for the next occasions of video packets.  Without considering any jitter, the worst extra-latency is 13.33 ms, if the 4 ms jitter is included, then the extra-latency in the DL alone is 17.33 ms.  For uplink, the audio (users 2 talks to user 1) can suffer from the same 17.33 ms extra-latency.  For the audio codec, another 32 ms latency should be included (20 ms framing for encoder and 12 ms algorithm lookahead (encoder\_decoder))  Note such an analysis has not considered TDD split between DL and UL. With the agreed TDD split DDDUU, alignment latency for audio stream, data stream and video stream should be considered also.  To simplify the modelling effort, we can consider merging data stream and audio stream so in total we have one video stream and one data/audio stream for DL.  We have  **Proposal: For AR2 downlink, 2 flows are modelled: scene/video + data/voice** |
| MTK | We support HW’s proposal to first model I/P frame for video streams and a 2:1 size ratio for I/P frame can be a starting point. For example, for 30Mbps video DL traffic, we can have:   * I: mean = 41667 bytes (20Mbps), STD = 5450 bytes, max size = 60000 bytes, period = 16.67ms * **P**: mean = 20833 bytes (10Mbps), STD = 2725 bytes, max size = 30000 bytes, period = 16.67ms   Since video stream is the dominant stream in DL, voice can be optionally modelled using for example VoIP traffic model. |

UL Traffic Model

In this section, we are discussing UL traffic model. To facilitate our discussion, Figure 3 summarized UL traffic characteristics from SA4 study. Scene update, video, audio, and data stream are illustrated together as a single stream in the figure, because their latency requirements are much less stringent than the other stream, control information/pose update (please note that given relatively loose latency requirements, we may assume that packets can be aggregated into a single stream at least for simplicity of simulations). On the other hand, the control information or pose update arrives more frequently with a tighter E2E latency requirement, it may not be assumed that this stream is aggregated with others. Please note that SA4 input on some of parameters, e.g., packet size, bit rate, latency requirement, periodicity, etc. are also summarized in Figure 1, based on which RAN1 can develop traffic model and relevant parameters.

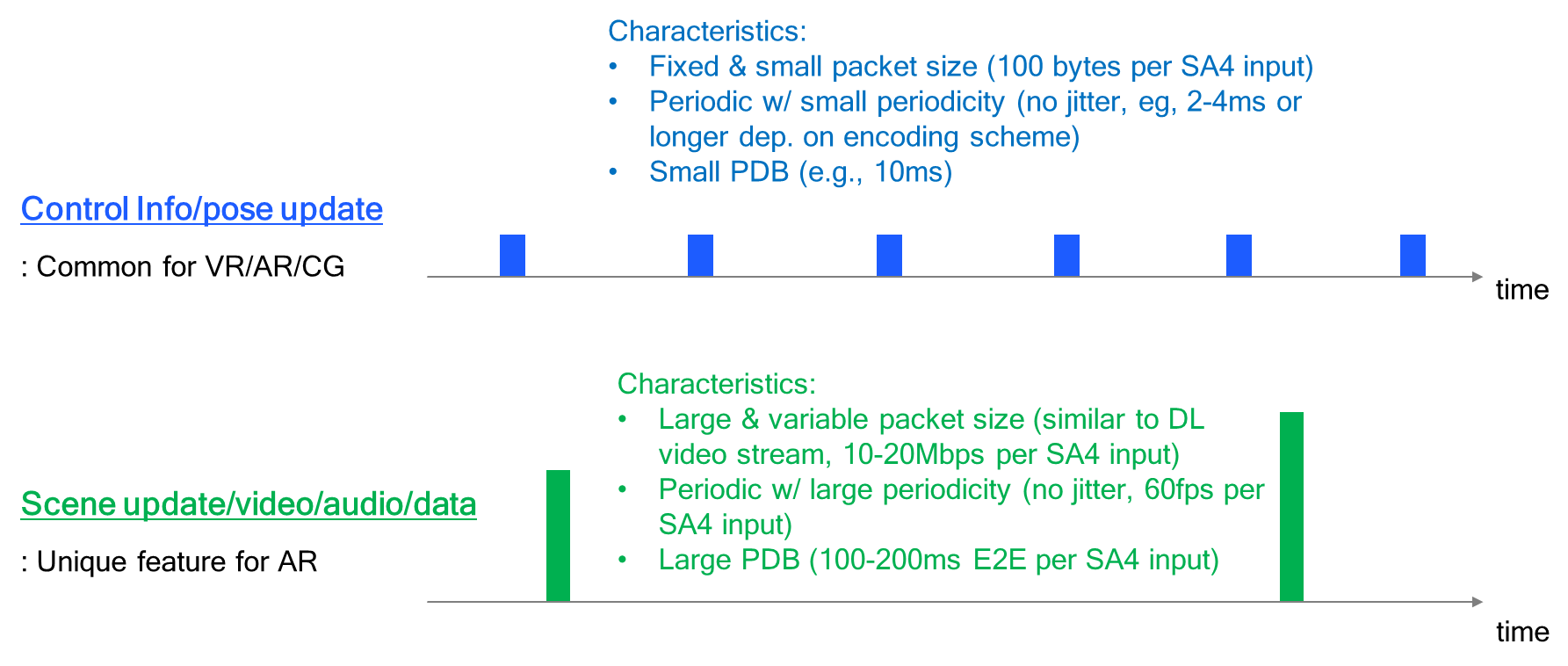


Figure 3. UL traffic characteristics per SA4 study

**# of streams in UL**

**Question 10.** Based on SA4 input summarized in Figure 3, please share your view on which/how many streams (e.g., 1 or 2 or more) should be explicitly modeled and evaluated for each application for UL. For example,

* CG/VR: single stream (pose/control)
* AR: dual flows (pose/control + scene/video/data/voice)

|  |  |
| --- | --- |
| Company | View |
| OPPO | We prefer one stream for each direction. |
| Ericsson | CG/VR/AR: single stream (pose/control) – but enough to ensure that UL is not limiting  AR: single stream – scene update (scene/video/data/voice)  Question: it is indeed stated 100-200ms E2E in the SA4 document, but it is also stated that the max slice latency is 80ms, which is the same as for CG. Can we clarify this? |
| CATT | Both single and multi-stream should be considered. |
| vivo | Support single stream for all application as starting point.   * CG/VR: single stream (pose/control) * AR: single stream (scene/video/data/voice) |
| Nokia, NSB | Agree with Ericsson proposal.  CG/VR/AR: single stream (pose/control)  AR: single stream – scene update (scene/video/data/voice) |
| Huawei, HiSilicon | For UL traffic of VR/CG, one single stream for pose/control information is reasonable. We are ok with this.  For UL traffic of AR, the video stream can be further modelled as I-frame stream and P-frame stream as we suggested in section 2.2.  Therefore, we have the following proposal:  *Proposal: For UL traffic of AR, the following options can be modeled and evaluated*   * *Option 1: 2 flows (pose/control + scene/video/data/voice)* * *Option 2: 3 flows (pose/control + I-frame of video + P-frame of video)* * *Option 3: 4 flows (pose/control + I-frame of video + P-frame of video + voice)* |
| Futurewei | We support single stream for each direction. |
| Sony | We can start with single stream for all applications |
| QC | Support the proposal. This model is simple yet capture different UL XR traffic characteristics well. In UL, two flows for AR is good enough to capture UL heavy traffic of AR. |
| LG | Single stream is preferred. |
| Apple | Downlink multiple flows should be modelled for AR2. It is clear SA4 includes multiple flows in the LS to RAN1. Moe specifically, audio stream, data stream and video stream are present in DL.  Assume video stream at 60 fps. The following timing diagram can be used to check the extra-latency experienced by the audio stream if audio packets wait for the next occasions of video packets.  Without considering any jitter, the worst extra-latency is 13.33 ms, if the 4 ms jitter is included, then the extra-latency in the DL alone is 17.33 ms.  For uplink, the audio (users 2 talks to user 1) can suffer from the same 17.33 ms extra-latency.  For the audio codec, another 32 ms latency should be included (20 ms framing for encoder and 12 ms algorithm lookahead (encoder\_decoder))  Note such an analysis has not considered TDD split between DL and UL. With the agreed TDD split DDDUU, alignment latency for audio stream, data stream and video stream should be considered also.  To simplify the modelling effort, we can consider merging data stream and audio stream so in total we have one video stream and one data/audio stream for DL.  On uplink multiple flow modeling:  The same analysis for downlink multiple flows applies to UL as well. To simplify the modelling effort, we can consider merging data stream and audio stream so in total we have 3 flows for uplink: one scene/video stream, one data/audio stream and one pose/control for UL.  We have  **Proposal: For AR2 uplink, 3 flows are modelled: scene/video + data/voice + pose/control** |
| MTK | We prefer to have 3 UL streams for AR as shown below. However, to facilitate progress, we can also live with FL’s proposal since it also seems reasonable.  [MTK preference]  For UL, we suggest to model three streams for AR:  1. [High priority] Gaming command traffic   * Interval: 8 ms * Size: 61 bytes 🡪 Bitrate: (7625 bytes/s)   2. [Normal priority] Background traffic   * Interval: 170 ms * Size: 360 bytes * (2140 bytes/s)   3. [Normal priority] Video traffic (1080x720, 30 fps)   * Interval: 33 ms * Size: 8k ~ 10k bytes   [Possibly also I/P for UL video model.] |

**Parameters for pose update/control information for UL**

Summary of 1st email discussion on parameters for pose update/control:

* Periodicity
  + 2/4/8/16 ms: OPPO, Interdigital, AT&T, Ericsson
  + 4 ms: MTK, Xiaomi, DCM, Sony
  + 2 & 16.7ms: QC
  + 2 & 4ms: vivo
  + 5 & 10ms: HW
* Packet size
  + 100 bytes: OPPO, QC, DCM, AT&T, vivo, Ericsson
  + 61 bytes: MTK
  + 100/250 byte: Interdigital
* Air interface PDB (latency requirement from UE MAC to gNB packet arrival)
  + 10ms: OPPO, MTK, Interdigital, AT&T, vivo, Ericsson

Based on SA4 input summarized in Figure 3 and based on companies views above, the following proposal is made.

**Proposal 11.** If pose update/control is evaluated as a stream in UL,

* Traffic parameters for Pose/control in UL
  + Periodic: 4ms (no jitter)
    - Other values can be optionally evaluated.
  + Fixed: 100 bytes (SA4 input)
  + PDB: 10 ms

**Question 11.** Please share your view on Proposal 11.

|  |  |
| --- | --- |
| Company | View |
| OPPO | Ok |
| Ericsson | Support |
| Xiaomi | Agreed. |
| CATT | OK. |
| Vivo | Fine with proposal 11. |
| Nokia, NSB | Agree |
| Futurewei | Support Proposal 11. |
| Sony | Support Proposal 11 |
| QC | Support |
| LG | Okay with Proposal 11. |
| InterDigital | Agree |
| MTK | Support |

**Parameters for scene upload/video/audio/data for UL**

Summary of 1st email discussion on parameters for UL scene upload/video:

* Bit rate:
  + 20 Mbps: QC, vivo, CATT, InterDigital
  + 10 Mbps: Ericsson (mandatory)
  + 2.4 Mbps: MTK
* Periodicity (or fps)
  + 30 fps: MTK
  + 60 fps (SA4 input): vivo, CATT, QC, InterDigital, DCM, Ericsson
  + 60 & 120 fps: OPPO, AT&T
  + 120 fps optional: DCM
* PDB
  + 40ms: MTK
  + 60ms: vivo, CATT, QC, InterDigital, Ericsson
  + 30/60/100ms: HW

Based on SA4 input summarized in Figure 3 and based on companies views above, the following proposal is made.

**Proposal 12.** If a single stream is evaluated for scene upload/video/audio/data,

* Traffic model for scene upload/video/audio/data
  + Periodic:
    - 60fps (SA4 input) – no jitter
    - Other values e.g., 30/120 can be optionally evaluated.
  + Data rate (MAC layer data rate)
    - 20 Mbps (Note: SA4 input is 10 or 20 Mbps)
    - Other values e.g., 2.4/10 Mbps can be optionally evaluated
  + Packet size: follows a probability distribution
    - As in DL, a Truncated Gaussian distribution is used.
      * Mean: derived from fps and data tare (same as in DL)
      * STD: Same as in DL
      * Max: Same as in DL
      * Min: Same as in DL
  + PDB:
    - 60 ms
    - Other values e.g., 30/40/100 ms can be optionally evaluated

**Question 12.** Please share your view on Proposal 12.

|  |  |
| --- | --- |
| Company | View |
| Ericsson | Support, with the exception of UL data rate – we should assume HEVC, i.e., 10Mbps  We reiterate our comment from the DL discussion that the packet size distribution should be symmetric to ensure that we get the desired data rate |
| CATT | OK |
| vivo | Fine with proposal 12. |
| Nokia, NSB | Support, in general. For the data rate, both 10 Mbit/s and 20 Mbit/s can be used. Regarding “same as the DL” for the packet size distribution, it is better to agree first on the DL distribution and then see if the same assumptions are applicable for the UL. |
| Huawei, HiSilicon | Similar to our comments above, suggest the following red changes:   * + Data rate ~~(MAC layer data rate)~~     - 20 Mbps @ 60fps (Note: SA4 input is 10 or 20 Mbps) |
| Futurewei | Support Proposal 12. |
| Sony | Support Proposal 12 |
| QC | We support the proposal. |
| LG | Okay with Proposal 12. |
| InterDigital | Generally ok with the proposal. We share similar view with Nokia. The parameters applied for the Truncated Gaussian distribution modelling the UL packet size distribution should be clarified when also clarifying the DL packet size distribution. |
| MTK | Support FL proposal |

Per UE KPI

The follow agreements on Per UE KPI were made in the last GTW session.

|  |
| --- |
| * Per UE KPI   + Baseline: A UE is declared a satisfied UE if more than X (%) of packets are successfully transmitted within a given air interface PDB. The exact value of X is FFS.   + FFS: In addition to the baseline, the following additional method is FFS     - When determining a XR/CG user is satisfied or not, the following factors are considered. FFS how to use those factors.       * Packet loss information       * Packet delay information       * Some XR/CG source related information if they can be available within RAN, e.g. the mapping between packet and slices or frames and the packet importance       * Multiple data streams traffic model       * FFS if there are multiple streams (if adopted) |

**X value**

**Question 13.** The following was agreed in RAN1, “Baseline: A UE is declared a satisfied UE if more than X (%) of packets are successfully transmitted within a given air interface PDB. The exact value of X is FFS.” Please share your view on the X value, e.g., X = 99 for all traffic, or X = 99 for scene upload/video/audio/data and X = ??? for pose update/control.

|  |  |
| --- | --- |
| Company | View |
| OPPO | We are fine with X=99 for all traffic. |
| Ericsson | 99% for all traffic |
| Xiaomi | Here is a question: If X=99, given fps = 60, there will be a frame error every several seconds. Will the user experience be too bad? |
| CATT | We are OK with the proposal |
| vivo | Support X=99 for all traffic as starting point. |
| Nokia, NSB | Support 99% for all traffic if majority agrees. |
| Huawei, HiSilicon | For video stream in a given XR or CG application, different packet error rates can result in different user experiences. To reflect different levels of user experience, more values can be evaluated, so the following X values are suggested.   * For Single video stream: X=99, 95 * For I-frame stream: X=99.9, 99.5 * For P-frame stream: X=99, 95 * Other values can be optionally evaluated |
| Futurewei | Support X=99 for all traffic. |
| Sony | We are OK to support X=99 for all traffic |
| QC | Support X=99 for all traffic. |
| LG | 99% for all traffic is preferred. |
| InterDigital | We share similar understanding with Huawei on how the different packet error rates experienced on different streams in an application could impact the user experience. We also think using X=99 for all traffic can be as a starting point. |
| MTK | We can not accept X=99 for all traffic. At least for UL gaming command/pose update, a more stringent requirement is required. We are proposing 99.9% for gaming command/pose update and 99% for other UL datas. For DL, We proposed X=99 for I frame and X=95 for P frame. |

**Other FFS points**

**Question 14.** As to the following FFS point, given that we have the baseline agreed already, FL proposal is to defer discussion on it to RAN1 105e. Please share your view on this. Regardless, please feel free to present details on the FFS point in the table below.

* FFS: In addition to the baseline, the following additional method is FFS
  + When determining a XR/CG user is satisfied or not, the following factors are considered. FFS how to use those factors.
    - Packet loss information
    - Packet delay information
    - Some XR/CG source related information if they can be available within RAN, e.g. the mapping between packet and slices or frames and the packet importance
    - Multiple data streams traffic model
    - FFS if there are multiple streams (if adopted)

|  |  |
| --- | --- |
| Company | View |
| OPPO | Ok |
| Ericsson | Strongly support |
| CATT | We are OK to discuss these issues at next meeting RAN1#104b-e but not RAN1#105-e |
| vivo | We support that additional metrics can be reported. We are also fine to further discuss the additional metrics for capacity evaluation in RAN1 #105. |
| Nokia, NSB | Agree |
| Huawei, HiSilicon | Considering the approaching deadline of RAN1#104-e, we are ok to defer the discussion to the next meeting, which is RAN1#104bis-e (not RAN1#105-e). |
| Futurewei | Ok. |
| Sony | Support the FL’s suggestion above. |
| QC | Agree with FL suggestion |
| InterDigital | Agree |
| MTK | Support |

# 1st Round Email Discussion

Traffic Model

### Statistical Traffic Model

As to XR traffic model for evaluation, RAN1 #103e made the following agreement.

|  |
| --- |
| Agreement (#103e)  Traffic model for DL and UL should reflect various aspects, e.g., various bit rates, variable frame/packet (definition of frame/packet to be clarified with traffic model as necessary) size, and periodicity (how to model jitter is FFS).  RAN1 will strive to conclude on detailed traffic models in the next RAN1 meeting (104-e) where SA4 outcome on traffic model is expected to be available.   * Statistical model is preferred. * It is preferred traffic model for both UL and DL have a certain degree of variability so that and the total number of traffic models can be reduced. * Note: Taking into account the fact that the decision on traffic models may hold many other crucial decisions, discussion on traffic model in the next RAN1 meeting is prioritized from the beginning. |

SA4 has been working on XR traffic model and delivered its outcome in a LS to RAN1[1]. Considering the SA4 input, companies discuss XR traffic model [2-20]. Table 2 captures the views from companies.

Table 2 Companies view on SA4 Traffic Models

|  |  |
| --- | --- |
| Company | View |
| FutureWei | However, from RAN1’s point of view, it will be more convenient if a statistical XR traffic model can be provided for system performance evaluation purpose.  *Proposal 1: FS\_NR\_XR\_eval adopts a statistical XR traffic model with inter packet arrival time modeled as a period plus some jitter, where the period is the inverse of frame rate, and with packet size modeled as Gaussian distributed.* |
| OPPO | Based on the outcome of XR/CG work in other groups (e.g., SA4) [3][4][5], we note there are too many configurations of XR/CG services. From our understanding, it is difficult for RAN1 to evaluate all the recommended configurations due to the huge workload. Thus, RAN1 should try to determine a limited number of traffic models based on the current 3GPP work on XR/CG and other groups’ inputs.  Proposal 9: RAN1 sends a LS to SA4 to provide a set of parameters for traffic models and ask for the corresponding values for each parameter, e.g., |
| Huawei | *Proposal 2: For the traffic model of XR and CG,*   * *Statistical model is adopted, and the statistical model can be developed based on SA4 outcomes.* * *RAN1 continues to discuss whether P-Trace based traffic model is applicable in RAN1 evaluations or not.* |
| Vivo | *Proposal 3: Support to adopt statistical model for XR evaluation as the starting point, P-Trace model can be optionally evaluated.* |
| ZTE | 1. RAN1 schedules a conference call for XR related QA sessions during the first week of RAN1#104-e, i.e. from January 25th covering at least the following discussion points   - The appropriate arrival rate for UL/DL traffic  - Whether/how to model jittering for UL traffic  - The IP packet level reliability requirement and corresponding slice recovery strategy that should be assumed in RAN simulation  - The E2E latency used for RAN1 evaluation  - The CRF setting and packet size assumption for the simulation  - Packet delay modeling  - whether to consider packet jittering in the reTx phase  - whether to consider jittering related to file size  - whether the packet delay should be i.i.d from a predefined distribution or fixed within a given buffer window.  - Clarification on typo regarding the delay threshold setting in cloud gaming  - Whether to prioritize some case of traffic modeling for both eyes |
| LG | *Observation 2: it may be necessary to communicate with SA4 on what kind of input RAN1 is expecting from SA4 and what kind of input SA4 can provide to RAN1 regarding XR traffic model.* |
| Sony | Proposal 3: RAN1 needs to derive traffic models and its parameters suitable for physical layer simulation based on higher-layer XR application parameters as identified by SA4. |
| Xiaomi | Proposal 1: Use the packet size distribution and packet inter-arrival time distribution achieved from SA4 trace data to randomly generate UE traffic in RAN1 evaluations. |
| Ericsson | 1. RAN1 should decide the exact video traffic parameters further based on input from SA WG4 XR study. The necessary parameters include a frame size in terms of mean, variance, the maximum and the minimum value at least for the minimal acceptable encoding rate, in addition to the frame generation interval. |
| QC | Proposal 1: RAN1 continues to assess SA4 traffic models and make conclusion at RAN1#105e.  Proposal 2: RAN1 supports the statistical DL traffic model in Table 1.  Proposal 6: RAN1 continues to discuss/determine the statistical models for VR1 and AR1 considering corresponding traffic model discussion in SA4. |
| Apple | **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms |   **We have the following in our contribution.**  **Observation 2: From SA4 traffic model on XR conversational, it is clear that uplink traffic is with substantial throughput requirements.**  **Observation 3: SA4 study on AR2 indicates multiple data flows are present in both downlink and uplink.**  **Proposal 1: It is key to include uplink traffic with substantial throughputs in the study of AR2.**  **Proposal 2: In RAN1 study, data flows with different QoS requirements in XR study should be modeled separately.**  **From SA4 study and LS to RAN1, there are two important points which should be reflected in RAN1 study:**   * + - 1. **Importance of UL traffic**       2. **Multiple data flows for each direction (DL and UL)** |

**Summary**

* Adopt a statistical model in RAN1 with parameters based on SA4 input: FutureWei, Huawei, vivo, QC, Sony, Xiaomi, Ericsson
* Communicate with SA4 to get further information (e.g., on model parameters) from SA4: Oppo, ZTE, LG

**Proposal 1**. RAN1 adopt a parameterized statistical traffic model for evaluation of XR and CG. RAN1 strive to agree on distributions and parameter values during RAN1 #104e, based on SA4 input.

**Question 1**. Please share your view on Proposal 1.

|  |  |
| --- | --- |
| Company | View |
| OPPO | We support statistical traffic model and agree to start work based on SA4 input. Once we find the information is missing for some parameter of the statistical traffic model, then we can ask SA4 for further information. In summary, we support FL’s proposal. |
| MTK | We are ok with Proposal 1 and no need for another LS to SA4 as this may take time and could be a long process.  We suggest RAN1 to use V-trace from SA4 to determine the statistical model (frame size distribution, mean, variance) with multiple DL data streams  For V-trace source file not available in the SA4 LS (Ex. AR, CG), RAN1 can determine the appropriate statistical numbers.  For instance, for V-trace of applications (AR, CG) not provided by SA4, adopt the following values for I frame:   * AR: mean = 83334 bytes (40Mbps), STD = 10900 bytes, max size = 120000 bytes, period = 16.67ms * CG: mean = 41667 bytes (20Mbps), STD = 5450 bytes, max size = 60000 bytes, period = 16.67ms   and the following values for P frame   * AR: mean = 41667 bytes (20Mbps), STD = 5450 bytes, max size = 60000 bytes, period = 16.67ms * CG: mean = 20833 bytes (10Mbps), STD = 2725 bytes, max size = 30000 bytes, period = 16.67ms |
| Xiaomi | We are fine with FL proposal. However, we suggest to send LS to SA4 on RAN1 decision to decide the distribution and parameter values based on SA4 input. If they have any concern, they can contact us. |
| QC | We agree on Proposal 1. |
| Nokia, NSB | We generally agree with Proposal 1. We just need to take into account that SA4 input might not be able to cover all parameters needed for RAN1. |
| Apple | **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms |   We are fine with statistical models, note SA4 study indicates multiple data flows are present for both DL and UL, the modeling of data flows other that of video stream is straight forward (constant rate is assumed). However, those data flows do have different periodicities than the video stream. For power consumption evaluation, if just the video stream is modelled then the evaluation may not be realistic, essentially with video stream only modeling, the UE can stay in a less-power consuming mode while there is no video packet, in reality the UE may need to be ready to receive/transmit packets for other data flows (audio, data stream, etc). We hope UE power consumption evaluation and capactiy evaluation are conducted with more realistic setup.  We are also aware of the modeling effort, so the number of data flows should be discussed, but a key fact from SA4 study should not be forgot in RAN1 study. |
| CATT | We are OK with the principle of Proposal 1. However, the XR traffic model by SA4 is an end-to-end traffic model between UE and XR server. For XR evaluation in RAN, the traffic arrival to gNB from XR server needs to incorporate the transport delay and jitter from XR server to the gNB. The traffic model in RAN1 is the queueing model at the gNB scheduler. Thus, a generic stochastic model (statistical data distribution and statistic packet inter-arrival time) is generally used with XR traffic model by SA4 as the reference for the parameter of XR traffic model. |
| Futurewei | We support Proposal 1 in principle. |
| InterDigital | We are ok with Proposal 1. For reducing the number of iterations for communicating with SA4 we think it would be worthwhile to consider deriving the statistical model based on the V-traces/P-traces provided by SA4 for the available use cases. |
| DOCOMO | We are fine with the FL proposal. |
| ZTE,Sanechips | Support. Statistical model is preferred. The traffic model should be based on SA4 input per 103-e agreement. RAN1 can communicate with SA4 to discuss details of traffic model. |
| AT&T | Agree with FL proposal |
| vivo | Agree with Proposal 1. The trace-based traffic model from SA4 can be further studied, e.g. P-trace or S-trace. |
| Huawei, HiSilicon | We support to adopt statistical model for RAN1 evaluation, and the distributions and parameters can be developed based on SA4 outcomes. As replied under Section 2.2.6, we think multiple data stream traffic model is essential for XR to accurately model real XR applications. So we suggest to add “of each data stream” in the proposal to be more accurate.  Meanwhile, RAN1 can continue to discuss whether P-Trace based traffic model is applicable in RAN1 evaluations or not. Currently, we think there is no need to send LS to SA4 to get further information.  The following changes in red are suggested:  **Proposal 1**. RAN1 adopt a parameterized statistical traffic model for evaluation of XR and CG. RAN1 strive to agree on distributions and parameter values of each data stream during RAN1 #104e, based on SA4 input. RAN1 can continue to discuss whether P-Trace based traffic model is applicable in RAN1 evaluations or not. |
| Sony | We support Proposal 1. |
| LG | Agree in principle with the proposal.  Considering the case where the packet data at the input of RAN simulator are packets over TCP/IP, we are not sure yet if we can directly derive a statistic model from video sequences or slices.  Anyway, we are supportive of trying to agree on the statistic model in this meeting and confirm or further refine it through communication with SA4. |
| Ericsson | Support |

### DL Traffic Model

In following sections, details of DL video traffic model are discussed.

#### Packet Arrival Modeling: Video Frame level vs IP-Packet level

In this section, we discuss DL *packet* arrival modeling. Depending on layer it models, the size of packet, traffic arrival pattern, performance requirement could be different. We capture two views on modeling: frame level model vs IP-packet level model, based on contributions [2-20]. In case of frame level, a *packet* corresponds to a XR video frame. Whereas, in IP-Packet level modeling, a *packet* corresponds to an IP-packet of which size is usually smaller (e.g., <=1500 bytes).

|  |  |
| --- | --- |
| Company | View |
| FutureWei | *Proposal 1: FS\_NR\_XR\_eval adopts a statistical XR traffic model with inter packet arrival time modelled as a period plus some jitter, where the period is the inverse of frame rate, and with packet size odelled as Gaussian distributed.* |
| OPPO | Proposal 4: Regarding the packet modelling for XR/CG, an application level packet is modelled as a packet during RAN1 simulation. |
| Huawei | *Proposal 4: For XR/CG performance evaluation, frame segmentation is not considered* *for simplicity, i.e., one video frame is modelled as one packet during simulation.*   * *Note: Each packet might be further segmented into one or multiple TBs for transmission in physical layer.* *The number of TBs and the size of each TB are up to radio resource, scheduling, etc.* |
| CATT | *Proposal 1: A data frame from Application is considered as the packet in the traffic model for XR. The jitter is characterized by the stochastic process of those segmented packets related to the same data frame.* |
| Vivo | *Proposal 5: For statistical model, following options for packet modelling can be considered,*   * *Option 1: an application level packet is modelled as a packet during simulation, i.e. one frame consisting of one or more IP level packets ≈ one packet in simulation.* * *Option 2: an IP level packet is modelled as a packet during simulation, i.e. one IP level packet ≈ one packet in simulation.* |
| InterDigital | * + Traffic file size distribution: [Truncated Gaussian distribution or Parero distribution with configurable mean, σ, min, max] (e.g. mean: 1200B, max: 1500B) |
| MTK | *Proposal 2: Two traffic models should be considered depending on the location of the XR/CG server (cloud/Edge).*  *Proposal 3: No MTU packet size restriction when the XR/CG server is located at the Edge.* |
| Samsung | For RAN1 evaluation purposes, option 1 is applicable – an application level packet can be modeled as a packet for SLS. |
| Ericsson | 1. When a jitter is modelled, it should be simple enough to simulate, e.g., uniform or truncated Gaussian, and the effect of IP segmentation needs to be avoided for RAN1 evaluation. |
| QC | * File arrival: a file is a burst of related data belong to a frame arriving together |
| Nokia | ***Proposal 3****: For VR1 packet size distribution, adopt a constant size packet in uplink and downlink. Assume a 1200 byte for the downlink packet size, while 100 byte for the uplink packet size.* |
| Apple | **Proposal 3:**   * **In the traffic model for XR, multiple data flows (e.g. for audio and video) for each direction (DL or UL) are generated for a UE;** * **Each data flows can be configured separately with** * **Periodicity** * **Packet size distribution (e.g. fixed or following a distribution)** * **Data flow specific latency and reliability requirements**   **Proposal 4:**   * **For XR conversational, for each data flow, the following can be configured separately:**   + **Periodicity**   + **Packet size distribution (e.g., fixed or following a distribution)**   + **Data flow specific latency and reliability requirements** * **downlink traffic includes the following data flows:**   + **2D video**   + **Audio**   + **Data stream** * **uplink traffic includes the following data flows:**   + **Video + Depth**   + **Front facing Camera**   + **3/6 DOF Pose**   + **Audio**   + **Data stream**   ***As indicated above, the number of data flows for DL and UL can be reduced to make modeling work easier. In SA4 study, multiple data flows are present for DL and UL, so we should reflect that in our study.***  ***For video stream, audio stream, and data stream, the periodicity can be different, e.g. video with 60 fps (50/3 ms for periodicity), audio is with 20 ms for periodicity, and data stream with 10 ms for periodicity.*** |
|  |  |

**Summary**

* (XR application video) Frame level modeling: FutureWei, Oppo, Huawei, CATT, vivo, MTK, Samsung, Ericsson, QC
* IP packet level modeling: InterDigital, vivo, MTK, Nokia

**Question 2**. Please present your view on frame level vs. IP packet level XR traffic modeling for evaluation. If you prefer IP packet level modeling, please explain how IP packets are mapped to a video frame, the relation of IP packet size to the corresponding video frame size, and how latency is evaluated (e.g., apply a given delay budget, e.g., 10ms for each IP packet, or latency is measured from the point when the first IP packet of a frame to the point when the last IP packet of the frame is successfully delivered to UE).

|  |  |
| --- | --- |
| Company | View |
| OPPO | Support Frame level modelling rather than IP packet level modelling. IP packed level modelling will involve more factors, which will make the modelling more complicated and make it more difficult for companies to convergence to the detailed models. |
| MTK | We prefer to use frame level modelling and define reliability/latency in terms of frame level. We understand that one frame still needs to be segmented into multiple packets in real network, but RAN1 can apply a simpler model in simulation, since in physical layer data are all transformed into TBs. |
| Xiaomi | Support frame level modelling rather than IP packet level. Another question for IP level simulation is whether we need to assume radio layer can get higher layer information such as whether two packets belong to the same frame, etc.? |
| QC | We prefer to have frame level modelling. In this model, by construction, data belonging to a frame has the same PDB. This is an important latency requirement to capture in XR evaluation. The periodic nature of packet arrival could be also easily captured.  In an IP packet level approach, it gets more complicated and difficult to capture these aspects. Model should be able to track a set of IP packets belonging to the same frame and they should have the same PDB. Two different levels of packet inter arrival time modelling would be needed; inter-frame and inter IP packet.  Thus, frame level modelling is preferred. |
| Nokia, NSB | We propose to model an IP packet. There is no need to go to the application level and model the application level frames for RAN1 simulations. The latency is measured for each IP packet, and satisfaction of a user is measured by monitoring how many of those IP packets are correctly received within the PDB. Adopting this approach, we don’t need to explicitly model video frames in RAN1 simulations. |
| Apple | We are flexible with the modelling choice here. It seems the difference between two approaches will be more pronounced, if the inter-IP packet generation is large. |
| CATT | Frame level should be sufficient for RAN1 XR evaluation in Uu interface. If IP level packet is defined, PDCH and RLC functions, such as segmentation/reassembly and RLC error control, needs to be included in the simulation. |
| Futurewei | We prefer frame level modelling. Using frame level modelling is sufficient for RAN1 evaluation purpose, just like in RAN1’s FTP traffic models, we do not model the IP level packets but use the file model instead. Using frame level modelling also make it easier to define KPI. |
| InterDigital | We are generally ok to consider either of the modelling options. Typically, when the frames are encoded into slices, each slice may be carried in different IP packets. The number of IP packets per frame/slice may also vary depending on the type of the frame (e.g. I, P, B). For simplifying RAN1 evaluations and avoiding any complexities related to segmentation and encoding, we think it fine to assume a packet corresponds to a frame. |
| DOCOMO | We prefer frame level modelling to relax the complexity of the models. |
| ZTE, Sanechips | To simplify RAN1 evaluation, it is proposed to use frame level modeling. |
| AT&T | Frame-based modelling can be a baseline working assumption. IP modelling is of course more realistic, however the impact on the results is not clear to justify the extra complexity and should be further evaluated. |
| vivo | Frame level XR traffic modelling is preferred for simplicity. |
| Huawei, HiSilicon | For simplicity, we suggest to adopt frame level modelling, i.e., one video frame is modelled as one packet during simulation. |
| Sony | Packet-level modelling may have some advantages but is not deemed necessary. Also frame-level simulation may be performed. |
| LG | IP packet level modelling is preferred. Being based on the higher layer data such as video sequences, slices may be easier to derive a statistic model, but the statistical properties do not maintain at the input of the RAN simulator, then it becomes incorrect anyway. Simplified statistic model based on the P-trace is preferred even if the statistical property does not exactly represent the input distribution of IP packets, we don’t think it is less representative than the frame level modelling. |
| Ericsson | Support frame level modelling. |

#### Traffic Inter Arrival Time Distribution

In this section, views on modeling of DL packet inter arrival time are captured.

|  |  |
| --- | --- |
| Company | View |
| FutureWei | *Proposal 1: FS\_NR\_XR\_eval adopts a statistical XR traffic model with inter packet arrival time modeled as a period plus some jitter, where the period is the inverse of frame rate, and with packet size modeled as Gaussian distributed.* |
| oppo | Proposal 2: For XR/CG evaluation, the data arrival is assumed as periodic and the periodicity can be two different options (X FPS is the frame rate):   * Option 1: Periodicity is 1/X s * Option 2: Periodicity is 1/2X s * X = 60, 120 |
| Huawei | *Proposal 6: For XR and CG performance evaluation, periodic traffic with frame arrival interval 1/FPS seconds is considered as a starting point.* |
| CATT | *Proposal 3:* *The jitter is modeled in packet interval time in the traffic model for XR, similar* *as the 3GPP RAN1 traffic models, e.g. FTP model 3.*  *Proposal 4: The Exponential distribution and Pareto distribution are both considered for the packet interval time under the different network load.* |
| vivo | * + Option 1: PDB is affected by jitter, e.g. residual PDB (k) = PDB – jitter (k), k is the index of a packet.   + Option 2: PDB is not affected by jitter, e.g. PDB is constant.   *Proposal 7: RAN1 should further consider and discuss the jitter modelling in traffic model.* |
| Sony | Proposal 4: RAN1 to decide whether network jitter in XR applications is considered or not in the evaluation assumptions. |
| MTK | *Proposal 4: Jitter modelling is required and shall be taken into account in simulations* |
| InterDigital | * + Traffic arrival distribution: [Quasi-periodic with configurable inter-packet arrival time duration] (e.g. FTP3, inter-packet arrival proportional to 1/frame-rate) |
| AT&T | Proposal 2: The XR Traffic model packet inter-arrival times should be based on a specified fixed interval (e.g. inverse of media frame generation rate). |
| Xiaomi | Proposal 1: Periodic traffic can be assumed for the DL and UL traffic of the VR service |
| Samsung | Proposal 2: For the traffic models for XR evaluations in RAN1, consider the following:   1. Periodic packet arrivals 2. Truncated Gaussian distribution for modeling packet size and jitter 3. Both DL and UL and discuss whether to prioritize UL 4. Conclude in RAN1#104-e on values of periodicity, packet sizes, target BLER, and end-to-end PHY latency, and on whether to concurrently simulate multiple data flows. |
| Ericsson | When a jitter is modelled, it should be simple enough to simulate, e.g., uniform or truncated Gaussian, and the effect of IP segmentation needs to be avoided for RAN1 evaluation. |
| Apple | * Each data flows can be configured separately with * Periodicity * Packet size distribution (e.g. fixed or following a distribution) * Data flow specific latency and reliability requirements   ***For video stream, audio stream, and data stream, the periodicity can be different, e.g. video with 60 fps (50/3 ms for periodicity), audio is with 20 ms for periodicity, and data stream with 10 ms for periodicity.*** |
| QC | * File arrival time for   + Without jitter, periodic with periodicity of sec, i.e.,   for   * + With jitter, a random jitter is added to each expected file arrival time, i.e., , where is a random variable following truncated Gaussian distribution truncated between - and |
| Nokia, NSB | ***Proposal 1****: For VR1, consider FTP Model 3 as a downlink traffic model. Assume an average bitrate of 58 Mbit/s (Full HD) and 90 Mbit/s (4K video).*  ***Proposal 6:*** *For AR1 option 1 (augmented video traffic in downlink), consider the FTP Model 3 as a downlink traffic model. Assume an average bitrate of 58 Mbit/s and 90 Mbit/s for 1080p (Full HD) and 4K video quality, respectively.*  ***Proposal 7:*** *For AR1 option 2 (video objects in downlink), consider the Periodic Traffic as a downlink traffic model. Assume a traffic source generating objects of 10kbit and 10Mbit every 5 seconds.* |
|  |  |

**Summary**

Packet inter arrival time modeling

* Periodic with jitter: FutureWei, Oppo, Huawei, InterDigital, AT&T, Xiaomi, Samsung, Apple, QC, CATT, vivo, MTK, Samsung, Ericsson
  + Jitter distribution
    - Truncated Gaussian: Samsung, Ericsson, truncated Gaussian
    - Uniform: Ericsson
* Exponential: CATT, InterDigital, Nokia
* Pareto: CATT

**Proposal 4**.

* RAN1 adopt periodic arrival with jitter for DL packet arrival modeling for XR/CG applications
  + Periodicity: 16.67ms(=1/60fps).
* Jitter follows truncated Gaussian distribution.
  + Please present your view on the values of mean, variance, truncation bound.
* The jitter for a packet affects the packet delay budget (PDB) of the packet, i.e., positive jitter (late arrival) gives smaller PDB and negative jitter (early arrival) gives larger PDB for a given packet.

**Question 4**. Please share your view on Proposal 4.

|  |  |
| --- | --- |
| Company | View |
| OPPO | Support the proposal in principle. I suggest to add the periodicity corresponding to 120fps since the UE supporting 120fps will be the trend in the near feature. As this work is an study item, the specification will be done in the future releases. By considering 120fps, we can better future-proof the output |
| MTK | We are fine with Proposal 4 and suggest that jitter follows truncated Gaussian distribution with   * Mean: 0 * STD: 3 * Max absolute value: 7   If multiple DL streams are applied, all of them would have the same distribution. If multiple streams have the same time arrival, they would share the same jitter value. If multiple streams are staggered, then the jitter values should be highly correlated. We prefer to have the same time arrival for multiple streams and same jitter. |
| Xiaomi | Agree with FL proposal. We can generally define the periodicity as 1/fps, and discuss the value of fps as a different issue. From our point of view, multiple values of fps may need to be evaluated. Same or different fps values can be set for different XR applications. |
| QC | We support periodic packet arrival with jitter modelling. We propose to evaluate 60Fps.  The impact of jitter needs to be captured such that the jitter could affect the PDB of each packet as described in proposal 4. We support modelling jitter as a random variable with truncated Gaussian distribution with mean of 0, variance of [2]ms and truncation bound of [-4, 4]. Note that truncation bound and frame rate should be chosen such that jitter should not introduce out of order packet arrival. |
| Nokia, NSB | We support FTP Model 3 with Exponential inter-arrival packet time. The reason is that it is well capable of simulating the burst of packets arriving to the destination. Moreover, in order to parametrize the model, we just need one value (rate or mean) in contrast to Proposal 4, where four values needed to parametrize the arrival time. Therefore, we propose the following:  RAN1 adopt FTP model 3 for the downlink arrival modeling for XR/CG applications, where packets inter-arrival time follows Exponential distribution with parameter Y. The exact value of Y is FFS. |
| Apple | ***For video stream, audio stream, and data stream, the periodicity can be different, e.g. video with 60 fps (50/3 ms for periodicity), audio is with 20 ms for periodicity, and data stream with 10 ms for periodicity.*** |
| CATT | We don’t agree with proposal 4. The statistic model of inter-arrival time of XR traffic model characterized network transport delay and jitter should be have one stochastic model instead of deterministic periodicity and statistic model for network jitter.  The stochastic model of Poisson process is used in 3GPP as FTP-1/FTP-2 for single session and FTP-3 as multi-Sessions. We believe that FTP-3 with Poisson inter-arrival time (aggregate traffic of multiple sessions with each session having exponential inter-arrival time) is sufficient for periodic traffic. |
| Futurewei | We support Proposal 4 in principle. We support a frame rate of 60 fps as a baseline. Companies can optionally bring in results for other frame rate(s) if they want to. |
| InterDigital | We are generally ok with FL’s proposal. For the parameter values corresponding to periodicities and those of truncated Gaussian distribution for jitter, we prefer using the values derived from SA4 traces or provided by SA4 for each of the considered/available applications. |
| DOCOMO | We generally support the FL proposal. 60 fps is the baseline and 120 fps can be reported optionally considering CG applications. |
| ZTE, Sanechips | (1)Ok with periodicity is 16.67ms.  (2)The traffic model should be determined based on the output of SA4. According to [S4aV200627], jitter includes at least PreEncoding delay and Encoding delay if frame segmentation is not considered. Both PreEncoding delay and Encoding delay can be constant, Equally distributed, or truncated Gaussian distributed. We think the distribution of Jitter should be determined first in SA4 so that RAN1 discussion should roll out accordingly  (3)Two types of delay are provided as shown below[S4aV200634]. We should determine how to define the PDB across the two delays firstly. The detail of PDB can be find in section 2.2.4.   |  |  | | --- | --- | | 4.2.5 Information for RAN Simulation A total of N=16 users are provided. RAN simulations may be carried out for N’=1, …, N users using traces for different users starting with 1, 2, … . All P-Traces cover 1min.  Two packet loss configurations are applied: 1e-3 for no maxSize restriction and 1e-4 for maxSize 1500 byte restrictions.  The delay threshold that is used in receiver is provided to be 60ms. This means that packets later than 60ms rendering time are considered late losses. Packets need to be delivered within 13-45ms to be useful in the receiver.  ...   |  | | --- | | {  ...  "RANConfiguration": {  "PLR": ["1e-2","1e-3","1e-4"],  "Delay": ["10","20"]  },  ...  } | | |
| AT&T | Evaluate both 60fps and 120fps |
| vivo | Periodic arrival with no jitter can be the starting point.  Periodic arrival with jitter can be adopted for evaluation after finalizing the jitter modelling, considering the SA4’s input.  Based on current SA4’s input, a video frame is transformed into multiple slices through compression and encoding, then the multiple slices are transformed into IP-packets through packetization as shown in the figure below. Because it is difficult to define a reference time for periodic packet arrival, how to calculate the jitter for a frame/slice/IP-packet based on a pre-defined reference time needs further discussions.    If the jitter is defined as the first IP packet arrival time of a frame (D) to the last IP packet arrival time of a frame (E) for each frame, the CDF of jitter follows truncated Gaussian distribution as shown in below figure based on SA4’s input.  However, in our understanding, the jitter should be deduced from the rendering time (A) of a frame to the last IP packet arrival time of a frame (E) for each frame as shown in below figure based on SA4’s input. The positive jitter would result in decreasing for the PDB of a packet, while the negative jitter does not affect the PDB of a packet since the jitter may be eliminated by Jitter Buffer Management. |
| Huawei, HiSilicon | Periodic traffic without jitter needs to be considered as a starting point for performance evaluation. This is simple and informative as a baseline. Jitter can be additionally considered as optional.  To provide better user experience, the frame rate could be higher than 60 fps. This is also reflected in TR 26.928 (copied below, red part). So we suggest that 90 fps and 120 fps can be considered as optional.  The periodicity of each data stream may be different if multiple data steam traffic model is considered, e.g., video stream and audio stream, I-frame stream and P-frame stream, may have different periodicity. So multiple data steam traffic model in Section 2.2.6 should be discussed firstly, and then come back to the details of each data stream.  ==  (copied from TR 26.928)  4.4 XR Engines and Rendering  …  4.4.3 Real-time 3D Rendering  …  Typically, rendering needs to happen in real-time for video and interactive data. Rendering for interactive media, such as games and simulations, is calculated and displayed in real-time, at rates of approximately 20 to 120 frames per second. The primary goal is to achieve a desired level of quality at a desired minimum rendering speed. The impact of the frame rate for the rendering pipeline is discussed in details in clause 4.2.2. The rapid increase in computer processing power and in the number of new algorithms has allowed a progressively higher degree of realism even for real-time rendering. Real-time rendering is often based on rasterization and aided by the computer's GPU.  == |
| Sony | We generally fine with proposal 4. Except, we should consider the periodicity by using the frame rate as suggested by SA4 (30 and 60 fps). |
| LG | We have a similar view with Nokia, CATT, vivo in that the inter-arrival time and jitter should be based on the IP packets at the input of RAN simulator. No need to directly set the periodicity based on the frame rate. Periodic arrival with jitter itself is okay to us. |
| Ericsson | Support. Jitter should be small (<5ms), to only consider variations in encoding. Jitter due to transport network variations should be excluded. |

#### Frame/Packet Size distribution

XR/CG frame sizes are varying depending on video encoding schemes. The following table captures views on packet size modeling.

|  |  |
| --- | --- |
| Company | View |
| FutureWei | *Proposal 1: FS\_NR\_XR\_eval adopts a statistical XR traffic model with inter packet arrival time modeled as a period plus some jitter, where the period is the inverse of frame rate, and with packet size modeled as Gaussian distributed.* |
| Oppo | Proposal 8: For RAN1 evaluation of VR2/AR2/CG, use the truncated Gaussian distribution to model the packet size. |
| Huawei | *Proposal 8: For XR and CG performance evaluation, the frame size is modelled as truncated Gaussian distribution. FFS: mean and variance.* |
| CATT | *Proposal 2: The Gaussian and Pareto distribution could both be considered for modeling the XR packet size.* |
| vivo | Table 3. DL traffic models for XR   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Traffic model | Model 1  (50Mbps) | Model 2  (50Mbps) | Model 3  (100Mbps) | Model 4  (100Mbps) | | Packet size distribution | Truncated Gaussian distribution | | | | | Mean packet size (Bytes) | 104000 | 52000 | 208000 | 104000 | | STD of packet sizes (Bytes) | 13000 | 6500 | 26000 | 13000 | | Maximum packet size (Bytes) | 162500 | 81250 | 325000 | 162500 | | Minimum packet size (Bytes) | 67 | | | | | Packet arrival interval (ms) | 16.67 | 8.33 | 16.67 | 8.33 | | Packet delay budget (ms) | 10 | | | |   Table 4. DL traffic model for Cloud Gaming   |  |  | | --- | --- | | Traffic model | Model 5 (25Mbps) | | Packet size distribution | Truncated Gaussian distribution | | Mean packet size (Bytes) | 52000 | | STD of packet sizes (Bytes) | 6500 | | Maximum packet size (Bytes) | 81250 | | Minimum packet size (Bytes) | 67 | | Packet arrival interval (ms) | 16.67 | | Packet delay budget (ms) | 15 |   *Proposal 9: For DL, traffic models in Table 3 and Table 4 are considered as the starting point for XR and Cloud Gaming evaluation, respectively.* |
| ZTE | 1. Adopt the three-step methodology to derive the traffic models for Pareto distribution of file size. |
| InterDigital | * + Traffic file size distribution: [Truncated Gaussian distribution or Parero distribution with configurable mean, σ, min, max] (e.g. mean: 1200B, max: 1500B) |
| AT&T | Proposal 1: The XR Traffic model packet size should be based on a truncated Gaussian distribution with specified mean size, variance, min file size, and max file size. |
| Samsung | 1. Truncated Gaussian distribution for modeling packet size and jitter |
| Ericsson | 1. The frame size for the video traffic may include a variance, e.g., Gaussian distribution, in time to be more realistic. |
| QC | * File size: random size following truncated Gaussian distribution with   + mean   + standard deviation   + Truncation between 0 and |
| Nokia, NSB | ***Proposal 3****: For VR1 packet size distribution, adopt a constant size packet in uplink and downlink. Assume a 1200 byte for the downlink packet size, while 100 byte for the uplink packet size.*  ***Proposal 9:*** *For AR1 packet size distribution, adopt a constant size packet in uplink and downlink. Assume a 1200 byte for a packet size in uplink and downlink.*  ***Proposal 13:*** *For CG packet size distribution, adopt a constant size packet in uplink and downlink. Assume a 1200 byte as a downlink packet size while 100 byte as an uplink packet size.* |
| Apple | ***Besides video stream, constant bit rate can be assumed for audio: 10.24 kbps for a packet (50 packets per second);***  ***5 kbps for a data stream packet (100 packets per second)***  **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms | |

**Summary**

Distribution of DL packet size

* (Truncated) Gaussian: FutureWei, Oppo, Huawei, CATT, vivo, InterDigital, AT&T, Samsung, QC
* Pareto: CATT, InterDigital, ZTE
* Constant packet size: Nokia

**Proposal 6**.

* RAN1 uses truncated Gaussian distribution as DL packet size distribution of XR/CG applications.
  + Please present your view on mean, variance, truncation bound for different applications.

**Question 6**. Please share your view on Proposal 6.

|  |  |
| --- | --- |
| Company | View |
| OPPO | Support the propsoal |
| MTK | Agree truncated Gaussian distribution as DL packet (= video frame) size distribution for XR/CG applications with statistical values derived from V-trace provided by SA4.  For V-trace source file not available in the SA4 LS (Ex. AR, CG), adopt the following values for I frame:   * AR: mean = 83334 bytes (40Mbps), STD = 10900 bytes, max size = 120000 bytes, period = 16.67ms * CG: mean = 41667 bytes (20Mbps), STD = 5450 bytes, max size = 60000 bytes, period = 16.67ms   and the following values for P frame   * AR: mean = 41667 bytes (20Mbps), STD = 5450 bytes, max size = 60000 bytes, period = 16.67ms * CG: mean = 20833 bytes (10Mbps), STD = 2725 bytes, max size = 30000 bytes, period = 16.67ms |
| Xiaomi | We agree with FL proposal |
| QC | We support using truncated Gaussian with mean =   (byte), standard deviation ], maximum size where is the DL bitrate (in Mbps) considered and is frame rates (fps) in DL. |
| Nokia, NSB | From the analysis of CG based on Google Stadia platform, we observe that the packet size is almost always constant. Therefore, we propose to consider the constant packet size. |
| CATT | We agree with the proposal to use truncated Gaussian. The mean and variance of truncated Gaussian random variable for XR traffic should be determined in RAN1#104-e based on SA4 XR traffic model. |
| Futurewei | We support Proposal 6 in principle. Qualcomm’s suggestion on mean, variance, and truncation bound is a good starting point. |
| InterDigital | We support FL’s proposal. For the parameter values for the DL packet size distribution, we prefer using the values derived from SA4 traces or provided by SA4 for each of the available applications. |
| DOCOMO | We support the proposal and also agree with refer SA4 traffic models for the statistical values. |
| ZTE, Sanechips | It is early to discuss the parameters. RAN1 should first determine between whetherPareto distribution as captured in the SID or Gaussian distribution should be appropriate setting for packet generation in simulator, based on SA4 input(e.g., P-trace) firstly.  We also want to finish the determination of traffic model as soon as possible. We can accept to discuss parameters of Pareto distribution, as captured in the SID to speed up the discussion. |
| AT&T | Agree with QC proposal as starting point |
| vivo | We support FL’s Proposal 6.  The parameters for truncated Gaussian distribution can be derived from the traffic models provided by SA4, e.g. the packet size distribution is derived based on frame sizes where the size of a frame is the sum of sizes for all IP-packets associated to the frame.  VR2: 2 eye buffers at 2Kx2K at 60 FPS, 8bit.  The truncated Gaussian distribution characteristics can be derived based on above figure.   |  |  |  | | --- | --- | --- | | Traffic model | VR2  (left and right eye frame arrive in turn) | VR2  (left and right eye frame arrive at the same time) | | Packet size distribution | Truncated Gaussian distribution | Truncated Gaussian distribution | | Mean (Bytes) | 43652 | 87304 | | STD (Bytes) | 10637 | 21087 | | Minimum (Bytes) | 14667 | 29334 | | Maximum (Bytes) | 90735 | 177845 | | Packet arrival interval (ms) | 8.33 | 16.67 |   For other applications, i.e. VR1, AR1, AR2, CG, corresponding packet size distribution can also be derived based on SA4’s input. |
| Huawei, HiSilicon | Ok to uses truncated Gaussian distribution as DL packet size distribution of XR/CG applications.  The data rate, packet size of each data stream may be different if multiple data steam traffic model is considered, e.g., FOV stream and non-FOV stream, video stream and audio stream, I-frame stream and P-frame stream, may have different data rate. So multiple data steam traffic model in Section 2.2.6 should be discussed firstly, and then come back to the details of each data stream. |
| Sony | In principle, support truncated Gaussian distribution. Furthermore, I-frame and P-frame may have different data rate. We need to discuss how to model it (e.g individual model). |
| LG | Our observation on size of IP packets comprising a frame is constant most of the time which is similar to Nokia’s observation above. We think constant packet size may be more realistic in that case and therefore should be considered. |
| Ericsson | Support proposal 6.  A general solution to the parameters of the distribution would be:   * Mean determined by data rate * Std deviation: 30% of mean * Min: 0 * Max: 2\*mean   This approach could be used for all periodic traffic models, irrespective of data rate. |

#### Packet Delay Budget

|  |  |
| --- | --- |
| Company | View |
| Vivo | *Proposal 10: The PDB requirement is assumed to be 10ms for XR traffic, and it could be relaxed to 15ms for Cloud Gaming traffic.* |
| ZTE | As discussed above, a packet which is delayed more than PDB is counted as a loss. The PDB is associated with two values, one is a fixed ‘Delay for RAN1’ (i.e., 20ms), the other is ‘remaining PDB for RAN1’ (i.e., 13-45ms for split rendering). The PDB can be determined as a minimum of Delay for RAN1 and remaining PDB for RAN1.   1. A packet which is delayed more than PDB is counted as a loss. The PDB is obtained by min(Delay for RAN1, remaining PDB for RAN1). |
| MTK | Proposal 9: For XR/CG evaluation, if one frame is mapped to one packet, a UE is satisfied if 99% files are delivered within packet delay budget (PDB) = 10ms. For requirements defined in SA2 LS S2-2009227 (5ms/99.9% and 10ms/99.9%), they can serve as optional (advanced) user satisfaction criterion. |
| **QC** |  |
| **Apple** | **The delay budgets for different streams are different, we can consider reduction of number of data flows**  **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms | |

**Summary**

Suggested packet delay budget values

* VR, AR: 10ms, 20ms
* CG: 15ms, 30ms

**Proposal 7**. RAN1 adopt over the air packet delay budget for packet transfer. The considered values for PDB are

* VR, AR: [10ms, 20ms]
* CG: [15ms, 30ms]

**Question 7**. Please share your view on Proposal 7.

|  |  |
| --- | --- |
| Company | View |
| MTK | We prefer to have one value for each application to limit simulation effort.   * VR, AR: 10ms * CG: 15ms |
| Xiaomi | We prefer to the same PDB for all applications, e.g. 10ms. |
| QC | We support proposal 7. PDB of 10ms and 20ms for VR and AR is a reasonable choice to evaluate AR/VR applications requiring different latency requirements.. PDB of 15ms and 30ms is a reasonable choice for CG as CG generally has less stringent latency requirement than AR/VR. |
| Nokia, NSB | We suggest to modify Proposal 7 to decrease the number of simulations:  **Proposal 7**. RAN1 adopt over the air packet delay budget for packet transfer. The considered values for PDB are   * VR, AR: [10ms, 20ms] * CG: Mandatory: [15ms]; Optional: [30ms] |
| Apple | According to SA4 study, the delay budget can be different for different data flows:  **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms | |
| CATT | We are OK of Proposal 7 as working assumption. The final budget should consider the network delay, inter-arrival time of actual traffic model and SA traffic model. |
| Futurewei | We support Proposal 7 in principle. To reduce the number of simulations, we propose using a single PDB of 15 ms for VR/AR/CG. The value of 15 ms is a good choice as it is in the middle of 10 ms and 20 ms for VR/AR and is part of the original PDB for CG. |
| InterDigital | We are ok with the values proposed by FL. However, to capture the application requirements more realistically and for reducing the number of simulation parameters the use of more stringent value for PDB per application, as proposed by MTK and Nokia, can be considered. |
| DOCOMO | We prefer the following:   * VR, AR: 10ms * CG: 15ms   20 ms for VR/AR and 30 ms can be optional. |
| ZTE, Sanechips | 1. In S4aV200634, **delay threshold and delay for RAN** are provided.   The packets scheduled later than delay threshold are considered as late losses. The transmission of packets needs to be limited within the two types of delay.  Delay of the packet arriving at the gNB/Radio is 15-47ms, and the remaining delay budget for RAN1 can be obtained by using delay threshold subtracting latency of the packet arriving at the gNB/Radio as shown in Figure 1.    Figure 1 Delay for Split Rendering  The PDB is associated with two values, one is a fixed ‘Delay for RAN’ (i.e., 20ms), the other is ‘remaining PDB for RAN1’ (i.e., 13-45ms for split rendering). The PDB should be obtained by min(Delay for RAN1, remaining PDB for RAN1).  (2)In S4aV200634, delay for RAN is [10ms, 20ms] for VR2, AR2 and CG. We prefer to follow SA4 output. |
| AT&T | Agree with Nokia |
| vivo | Support Proposal 7 in principle. To reduce simulation burden, suggest to narrow down the combination, e.g.   * VR 10ms * AR 20ms * CG 15ms |
| Huawei, HiSilicon | For a given XR or CG application, different roundtrip interaction delays can result in different user experiences. To reflect different levels of user experience, more values can be evaluated, so the following red values are suggested.   * VR/AR: [5ms, 10ms, 20ms, 30ms] * CG: [15ms, 30ms, 50ms]   Meanwhile, the PDB of each data stream may be different if multiple data steam traffic model is considered, e.g., video stream and audio stream may have different PDB. So multiple data steam traffic model in Section 2.2.6 should be discussed firstly, and then come back to the details of each data stream. |
| Sony | “air packet delay budget” might be too vague or at least, can we provide the definition of “air packet delay budget”?  We are fine with those number, at least as a starting point (e.g. keep [] for now). |
| LG | Single representative value is preferred. For the single value, 10 ms should be considered. 10 ms for VR/AR and 15 ms for CG is also fine taking into account the fact that the latency can be a bit relaxed for CG. |
| Ericsson | Support. One value for AR/VR (10ms) and one for CG (15ms) is preferred. |

#### Bitrates for evaluation

|  |  |
| --- | --- |
| Company | View |
| Oppo | Proposal 5: For VR2/AR2/CG evaluations, consider the following typical DL rates   * 90Mbps * 45Mbps * 24Mbps * 12Mbps. |
| Huawei | *Observation 1: For 4K VR video at 60fps, the* *bitrate would be about 60 Mbps. For CG, the bitrate would be about 35Mbps.* |
| CATT | *Observations 1: For the basic 4K VR video, the average bit rate would be 63Mbps per eye for the independent VR video streaming and 116Mbps for dual-eyes streaming. The new VVC codec would have addition 30-50% better compression ratio.* |
| MTK | Table 3: Cloud gaming traffic parameters |
| QC |  |
| Nokia | ***Proposal 1****: For VR1, consider FTP Model 3 as a downlink traffic model. Assume an average bitrate of 58 Mbit/s (Full HD) and 90 Mbit/s (4K video).*  ***Proposal 6:*** *For AR1 option 1 (augmented video traffic in downlink), consider the FTP Model 3 as a downlink traffic model. Assume an average bitrate of 58 Mbit/s and 90 Mbit/s for 1080p (Full HD) and 4K video quality, respectively.*  ***Proposal 7:*** *For AR1 option 2 (video objects in downlink), consider the Periodic Traffic as a downlink traffic model. Assume a traffic source generating objects of 10kbit and 10Mbit every 5 seconds.*  ***Proposal 11:*** *For CG, consider the FTP Model 3 as a downlink traffic model. Assume a traffic source generating 29 Mbit/s, and 45 Mbit/s as bitrate for 1080p (Full HD) and 4K video quality, respectively.* |
| Apple | ***Constant date rate for audio and data stream should be considered.*** |

**Summary**

* VR: 12, 24, 45, 90, 60, 116, 30, 60 Mbps
* AR: 12, 24, 45, 90, 30, 60 Mbps
* CG: 12, 24, 45, 90, 35, 5-20, 10-30, 8, 30 Mbps

Note that SA4 has provided following set of DL bit rates.

* VR2: 30, 60 Mbps
* AR2: 30, 60 Mbps
* CG: 8, 30 Mbps

**Proposal 8**. RAN1 adopt following DL bit rates based on SA4 input.

* AR/VR: [30, 60] Mbps
* CG: [8, 30] Mbps

**Question 8**. Please share your view on Proposal 8.

|  |  |
| --- | --- |
| Company | View |
| OPPO | It seems the values in Proposal 8 are the media bitrates suggested by SA4. However, in S4aV200632 or S4aV200640 (included in the SA4 LS), the bitrate in RAN should be 1.5 times of media bitrate. Thus, the above values should be changed to 30\*1.5, 60\*1.5, 8\*1.5, 30\*1.5. I copied the corresponding part as below |
| MTK | We prefer to have one value for each application.   * AR/VR: 60 Mbps * CG: 30 Mbps |
| Xiaomi | We also prefer to a single value for each application, 30 for CG and 60 for AR/VR.. |
| QC | We support proposal 8. The values are reasonable choices to evaluate applications requiring different bitrate applications. |
| Nokia, NSB | In general, we support Proposal 8. However, we think the lowest values for AR/VR and CG refer to the resolution 720p, which might be a bit low for the considered applications. We also think it is beneficial to consider a 4K resolution and the associated bit rates at least for CG. We thus propose to modify the proposal 8:  **Proposal 8**. RAN1 adopt following DL bit rates based on SA4 input.   * AR/VR: [~~30~~, 60] Mbps   CG: [~~8~~, 30, 45] Mbps |
| Apple | As we discussed above, XR traffic is not only about video stream, other streams should be considered also.  **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms | |
| CATT | We are OK with proposal 8 as working assumption pending the consistence with SA4 traffic model. |
| Futurewei | We prefer to have a single value for each application to reduce the number of simulations. We suggest 60 Mbps for AR/VR and 30 Mbps for CG, similar to what MTK and Xiaomi suggested. |
| InterDigital | We have a similar understanding with Nokia regarding bitrates for 4K resolution. As such, the proposed values for DL bitrates should be revised to 60Mbps for AR/VR and [30,45] Mbps for CG. |
| DOCOMO | We also prefer to have a single value for each application as follows:   * AR/VR: 60 Mbps * CG: 30 Mbps |
| ZTE, Sanechips | According to SA4 input, we prefer to include 45Mbps for both AR/VR and CG.   |  | | --- | | S4aV200627  4.1.3 Packet Generation  For each of the users, packet traces are generated. Two configurations are provided  1500 byte max packet size (addressing the cloud server case in S4aV200607)  Unlimited packet size, i.e. each slice results in a packet (addressing the edge serve case in S4aV200607)  The bitrate is assumed 45 Mbit/s. For a 30 Mbit/s bitrate, this would allow an excess of 1.5. This aligns with the parameters in S4aV200607. | |
| AT&T | Agree with Nokia |
| vivo | Support Proposal 8 in principle. To reduce simulation burden, suggest to narrow down the combination, e.g.   * AR/VR 60Mbps * CG 30Mbps |
| Huawei, HiSilicon | The data rate of each data stream may be different if multiple data steam traffic model is considered, e.g., FOV stream and non-FOV stream, video stream and audio stream, I-frame stream and P-frame stream, may have different data rate. So multiple data steam traffic model in Section 2.2.6 should be discussed firstly, and then come back to the details of each data stream. |
| Sony | Support Proposal 8. |
| LG | We also prefer a single value per each case. The suggestion from a few companies (60Mbps for AR/VR and 30 Mbps for CS) works for us. |
| Ericsson | Support. Lower values can be mandatory. Modern video codes tend to produce lower rates. |

#### Multiple Flows

|  |  |
| --- | --- |
| Company | View |
| Huawei | *Proposal 11: The following general traffic model is considered for the XR and CG:*   * *#M data streams for DL and #N data streams for UL, where each data stream has separate* |
| InterDigital | * + Number of data streams: [Configurable number of streams, configurable traffic parameters common to all streams] (e.g. isochronous multi-stream with bounded latency) |
| MTK | *Proposal 5: traffic model shall take into account different traffic types and possibly differentiated frames within the same application, in both UL and DL directions* |
| Apple | Proposal 2: In RAN1 study, data flows with different QoS requirements in XR study should be modeled separately.  **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms | |

**Summary**

* Support multiple traffic flows in DL/UL with different QoS requirements/date rates: Huawei, InterDigital, MTK, Apple

**Question 9**. Some companies are proposing to explicitly evaluate multiple traffic flows in DL/UL with different QoS requirements/date rates. Please share your view on this. If you support, please explain the details of how to evaluate multiple traffic flows, via examples, e.g., voice flow of X1 Mbps, Y1 fps, Z1 delay budget and video frame of X2 Mbps, Y2 fps, Z2 delay budget, how to define whether UE is satisfied or not, etc.

|  |  |
| --- | --- |
| Company | View |
| OPPO | We support single traffic flow in the simulation. If there are multiple traffic flows with different QoS requirements, should the NW differentiate different flows and optimize the scheduling? The performance will heavily depend on the implementation of QoS-aware scheduling. As a result, it will be difficult for a fair comparison. For example, if one result show better performance flow A and worse performance for flow B compared to another results, what’s the conclusion we can made based on these two results.  On the other hand, multiple traffic flows will require additional efforts for evaluation/simulation. |
| MTK | For DL, we suggest to model two streams:  1. DL stream 1: I frame in V-trace  2. DL stream 2: P frame in V-trace   * + FFS: Different QoS (PER and PDB) requirement for I/P frame   + FFS: FoV v.s. non-FoV streams   For UL, we suggest to model three streams for AR:  1. Gaming command traffic   * Interval: 8 ms * Size: 61 bytes   2. Background traffic   * Interval: 170 ms * Size: 360 bytes   3. Video traffic (1080x720, 30 fps)   * Interval: 33 ms * Size: 10k bytes   For UL, we suggest to model two streams for CG/VR:  1. Gaming command/pose information   * Interval: 4 ms * Size: 61 bytes   2. Background traffic   * Interval: 170 ms * Size: 360 bytes   FFS: Different QoS (PER and PDB) requirement for different UL streams |
| Xiaomi | We support to use a single traffic flow in the evaluation. For audio and data traffic, it has much lower data rate compared with video traffic. |
| QC | We think **single flow could be supported in DL**. In XR application, when video and voice are transmitted together, the amount of data traffic is usually much larger than voice traffic and performance requirement of video is tighter than that of voice in terms of bitrate, reliability, and latency. Thus, we think having additional modelling for the voice traffic would not make a significant impact to the result. Therefore, given the amount of complexity, we think, in DL, evaluating single flow would be good enough.  **In UL, up to two flows can be discussed for AR use case**– one for pose and the other for scene update, especially in case traffic arrival periodicity, delay/reliability requirement, and data rate requirement for these two traffics are very different; In this case, these two traffics could differently affect resource allocation/gNB scheduling, UE capacity and power, which is worth evaluating.  For UL pose, we can discuss [0.2, 1] Mbps, 60 - a few hundreds fps, PDB=[10]ms. [PER=1e-2]  For UL scene upload, we suggest to use [10, 20] Mbps, [60]fps, PDB=[100]ms. [PER=1e-2]  UE is satisfied if both flows satisfied reliability and latency requirement. |
| Nokia, NSB | We do not support modelling multiple flows per each direction. We propose to focus on one flow (e.g., most challenging one) per each direction as a baseline. |
| Apple | **For AR2, it is clear multiple data flows are present for both DL & UL. Study with a single flow won’t be realistic. Considering evaluation effort, we are open to discussion in reducing the number of data flows for RAN1 study.**  **Note the LS from SA4 to RAN1 includes details for traffic modeling which can be found at**  **https://www.3gpp.org/ftp/tsg\_sa/WG4\_CODEC/3GPP\_SA4\_AHOC\_MTGs/SA4\_VIDEO/Docs/S4aV200640.zip**  **One table is pasted below:**   |  |  |  | | --- | --- | --- | | **Media** | **Format and Model** | **E2E Latency requirement** | | 3/6DOF Pose | Same as for split rendering | UL: 5-10 ms | | Video + Depth | 1080p, Capped VBR 10/20 Mbit/s for UL | Conversational 100ms, 200ms | | 2D Video is split rendering | 1080p or 4K (2 eyes) same model as split rendering | 60ms  100ms | | Front Facing Camera\* | 720p, CBR 3 Mbit/s for UL | Conversational  100ms, 200ms | | Audio (MPEG-H) | 256/512 kbps for both UL/DL | Conversational 100ms, 200ms | | Data Stream | 0.5 Mbps for both UL/DL | Conversational 100ms, 200ms | |
| CATT | **We are Ok to model multi-flow with FTP-3. FTP-3 with Poisson inter-arrival distribution is a multi-flow with each flow as exponential interarrival.** |
| Futurewei | We support modelling single traffic flow on each direction as a baseline. Companies can optionally bring in evaluation results for multiple traffic flows if they want to. |
| InterDigital | We think it is important to model different number of flows in DL and UL per application, where each flow may be subject to different QoS requirements (e.g. PDB, PER), for realistically reflecting the XR/CG application and to capture the impacts of potential QoS degradation of a subset of flows on overall user satisfaction. For minimizing the number of flows considered, the flows which have common traffic characteristics (e.g. periodicity) may be grouped and evaluated collectively with respect to the QoS requirements. |
| DOCOMO | We support to use single traffic flow for DL in the evaluation, while it is obvious that multi-flow is needed for UL. |
| ZTE, Sanechips | (1)DL may need consider two streams for two eyes separately.  (2)UE is satisfied if all streams are satisfied. |
| AT&T | Modelling only a single traffic flow in each direction is not realistic. At the same time as a compromise to minimize simulation complexity we are OK with a single DL flow but multiple UL flows as proposed by QC. |
| vivo | Support to adopt single traffic flow for DL/UL evaluation as baseline.  FFS multiple flows, e.g. how to model multiple flows and how to determine the requirements for different flows, what are the traffic characteristics for different flows. |
| Huawei, HiSilicon | For a given XR or CG application, there can be multiple data streams in DL/UL. And each data stream may have different traffic characteristics (e.g., periodicity, data rate, packet size, etc.) and QoS requirements in DL/UL. For example,   * FoV stream and non-FoV stream: the data rates of the two streams are different since high/low resolution videos are conveyed on each stream (see SA4 study outcome in S4aV200632 below, red part). * I-frame and P-frame: Different frame types may be of different size and importance. * Video stream and audio stream: Video and audio can have different periodicity, data rate, latency requirement, etc. * UL pose/control and UL scene update: The data rate, periodicity, latency requirement are very different. * Etc.  |  | | --- | | (Copied from S4aV200632)  …  6 VR1: “Viewport dependent streaming”  …   * + 1. Output traffic characteristics * Data rate range:   + per tiled streaming: 0.71~1.43 Mbps   + FoV Area Streaming: (0.71~1.43)\*18 Mbps   + low-resolution 4K omnidirectional streaming: 6-8Mbps |   If RAN1 only considers one data stream, it seems there is almost no difference between XR and URLLC traffic, and RAN1 is just going to evaluate URLLC traffic with more challenging requirements instead of XR traffic. Such evaluation results cannot reflect real XR applications and is not so meaningful to gain insight on aspects like how well NR network can support XR services, what’s the dominating factoring on supporting XR services, etc.  As agreed in RAN1#103-e that *“It is preferred traffic model for both UL and DL have a certain degree of variability so that the total number of traffic models can be reduced”*, a general model for all the five applications are preferred. The differences among each application could be the detailed values, which can be further discussed and figured out based on SA4’s outcomes.  So we suggest to agree the following proposal. And since we still have one more week, RAN1 can strive to agree on the detailed parameters during RAN1#104e, based on SA4 input.  *Proposal: The following general traffic model is considered for the XR and CG:*   * *#M data streams for DL and #N data streams for UL, where each data stream has separate* * ***Packet size distribution*** * ***Packet arrival interval*** * ***QoS requirement*** * *RAN1 strive to agree on the above parameters during RAN1#104e, based on SA4 input.* |
| LG | A single traffic flow per each link direction is preferred. Having 2 UL flows as optional can be further discussed, but the conclusion can only be made after the traffic model is agreed. |
| Ericsson | Do not support. It will be unfeasible to agree on criteria, and evaluation comparisons will be impossible |

#### Other issues

|  |  |
| --- | --- |
| Company | View |
| vivo | *Proposal 6: Adopt random offset for modelling traffic arrival offset among UEs per cell.*  *Proposal 12: For statistical XR traffic model, the following* *two traffic source types can be considered for evaluation, assuming frame rate is X FPS.*   * *Traffic source type 1: every 1/X s, the packets of both eyes arrive at the same time for each frame.* * *Traffic source type 2: every 1/(2\*X) s, the packets of left eye and right eye arrive in turn, e.g. the packet of left eye arrives at odd frames, while the packet of right eye arrives at even frames.*   *Proposal 13: For DL/UL packet arrival time relationship, following options can be considered,*   * *Option 1: DL and UL packet arrival times are independently modelled.* * *Option 2: DL and UL packet arrival times are correlated, e.g.* *the uplink pose information triggers the corresponding downlink traffic, where the interval between DL and UL packet arrival times may be uniform or subject to a certain distribution.* |
| ZTE | According to [7], traffic modeling for both eyes could be considered for the case of split rendering. In this case, we believe eye staggering should be considered and modeled since packetization based on smaller packets could be used to model the case of packets for both eyes arriving at the same time. |
| Ericsson | 1. Traffic arrival time offset among XR users needs to be included, e.g., random offset with the simple uniform distribution of [0 1/FPS] where FPS is a frame refresh rate |
| QC | Traffic start offset: different users could have different traffic start offset, which shifts the actual file arrival times of each UE |
| InterDigital | For the evaluations of capacity, both random offset and uniform offset can be considered. |

**Summary**

* Interleaved two eye buffer modeling: vivo, ZTE, QC
* Traffic arrival time offset across UEs: vivo, Ericsson, QC
* Dependency of DL and UL traffic: vivo

**Question 10**. Please share your view on (i) whether/how to evaluate two eye buffers, (ii) how to model traffic arrival time offset across UEs, (iii) whether/how to evaluate dependency of DL and UL traffic.

|  |  |
| --- | --- |
| Company | View |
| OPPO | 1. We support to consider the case of interleaved two eye buffer model. We prefer to reuse the above modelling except that the bitrate is 50% of that in section 2.2.5  2. We prefer to evaluate DL and UL independently. However, we also can live with the joint evaluation of DL and UL if majority companies support it. |
| MTK | (i) No need to evaluate two eye buffers because we do not see the need in RAN1 unless the two buffers are staggered and this could be possibly useful for better capacity. Also, V-trace provided by SA4 does not have L/R information.  (ii) uniform random offset from 0~16.67ms  (iii) Interesting idea but how to model this dependency is FFS |
| Xiaomi | We support to use separate evaluation for UL and DL. |
| QC | It may be beneficial to evaluate **two eye buffers** as an optional configuration as there may be diverse implementations to handle two eye buffers in reality. We can evaluate and compare performance among different options for packet arrival and PBD, e.g., interleaved arrival of two eye buffers at gNB vs. aligned arrival of two eye buffers at gNB, where how PDB is defined can be further discussed.  **As to Traffic arrival time across UE,** random offset among UEs can be the baseline, while we can evaluate the performance impact of equal or different offsets among UEs (assuming the offsets can be coordinated between gNB and edge server).  **Evaluating dependency between UL and DL** requires knowledge and modelling in network side, e.g., rendering time, network delay, etc. At this point, it seems that there is no strong reason to capture this. Having no dependency could also make it easy to do independent DL or UL evaluation. |
| Nokia, NSB | Our view is as following:  (i) whether/how to evaluate two eye buffers - the packets of both eyes arrive at the same time for each frame.  (ii) how to model traffic arrival time offset across UEs – uniform time distribution.  (iii) whether/how to evaluate dependency of DL and UL traffic - an independent modelling of DL and UL traffic as a baseline is preferred. |
| CATT | Two-eye buffering is one of the use cases for multi-session Poisson inter-arrival if we use FTP-3. |
| Futurewei | Our views are as follows:   1. No need to model two eye buffers. 2. Traffic arrival time offset across UEs can be modelled as uniform distribution. 3. DL and UL should be modelled independently to reduce the complexity of the evaluation work. |
| InterDigital | (i) We prefer the interleaved model where the frames associated with 2 eye buffers can be staggered to reduce the total bitrate. Since the P-traces contain the information on left/right eye, it would be possible to evaluate the performance of each buffer separately and collectively.  (ii) Uniform distribution for modelling the packet arrival offset among UEs  (iii) Independent arrival for UL and DL is preferred for simplifying the evaluation effort |
| DOCOMO | (i) can be an optional evaluation  (ii) uniform distribution  (iii) prefer independent modelling of DL and UL. |
| ZTE, Sanechips | 1. Eye staggering should be considered and modeled where the packet size is smaller. 2. Random distribution, which is more practical from application perspective, is preferred for traffic arrival time offset across UEs considering jittering effect. 3. We prefer to evaluate DL and UL independently. Some UL DL correlation may exist in aspects such as latency which could be considered as PDB requirement to the simulation. |
| AT&T | * Traffic arrival time offset across UEs: Ok with uniform offset as baseline but companies can simulate other distributions optionally if they believe there are important insights in that case * Dependency of DL and UL traffic: Perhaps this can be captured as part of the KPIs. The DL and UL flows are independent, however some per-user statistics which aggregate across DL/UL may be evaluated in addition to the separately collected DL and UL KPIs. |
| vivo | Regarding two eye buffer modelling, both the packets of two eyes arrive at the same time and interleaved two eye arrival should be considered.  With respect to traffic arrival time offset across UEs in a same cell, random offset may be considered for simplicity.  For dependency of DL and UL traffic, separate evaluation for DL and UL traffic is preferred to avoid complicated modelling and inefficient simulation. |
| Huawei, HiSilicon | (i) On two eye buffer modelling: simultaneous arrival can be considered as the baseline for simplicity.  (ii) On traffic arrival time offset across UEs: random offset can be considered, e.g. uniform distribution.  (iii) On dependency of DL and UL traffic: suggest no dependency. Application layer will use UL pose/control to render frame, but this is transparent to RAN transmission. So from RAN’s perspective, both DL and UL are periodic traffic, and there is no relationship between them. There is no need to model the interaction between UL and DL in RAN1. |
| Sony | We support   * Adopt random offset for modelling traffic arrival offset among UEs per cell (Proposal 6 – Vivo) ;and   DL and UL packet arrival times are independently modelled (Proposal 13 opt1 – Vivo) |
| Ericsson | 1. Do not support – too complicated. 2. Randomly distributed traffic arrival across UEs 3. Too complicated |

### UL Traffic Model

This section discusses UL traffic model.

Table 6 captures the views on UL traffic model from different sources.

Table 6 Views on UL Traffic Model

|  |  |
| --- | --- |
| Source | View |
| Oppo | Proposal 6: For VR2/AR2/CG evaluations, RAN1 doesn’t need to consider the UL traffic and only evaluate the DL traffic in the NR system. |
| Vivo | Table 7. UL traffic model for media service   |  |  | | --- | --- | | Traffic model | Model 6 (13Mbps) | | Packet size distribution | Truncated Gaussian distribution | | Mean packet size (Bytes) | 27040 | | STD of packet sizes (Bytes) | 3380 | | Maximum packet size (Bytes) | 42250 | | Minimum packet size (Bytes) | 67 | | Packet arrival interval (ms) | 16.67 | | Packet delay budget (ms) | 10 |   *Proposal 14: For UL* *media services, the traffic model in Table 7 is considered as the starting point for the evaluation of AR case.*  Table 8. UL traffic model for interactive/pose services   |  |  | | --- | --- | | Traffic model | Model 1 (200kbit/s) | | Packet size distribution | Fixed, 100Bytes | | Packet arrival interval (ms) | 4 | | Packet delay budget (ms) | 10 |   *Proposal 15: For UL interactive/pose services, the traffic model in Table 7.* UL traffic model for media service   |  |  | | --- | --- | | Traffic model | Model 6 (13Mbps) | | Packet size distribution | Truncated Gaussian distribution | | Mean packet size (Bytes) | 27040 | | STD of packet sizes (Bytes) | 3380 | | Maximum packet size (Bytes) | 42250 | | Minimum packet size (Bytes) | 67 | | Packet arrival interval (ms) | 16.67 | | Packet delay budget (ms) | 10 |   *Proposal 14: For UL* *media services, the traffic model in Table 7 is considered as the starting point for the evaluation of AR case.*  Table 8 *is considered as the starting point for XR and Cloud Gaming evaluation.* |
| InterDigital | Proposal 3: The configurable parameters in the generalized XR traffic model for UL transmissions are:   * + Traffic arrival distribution: [Quasi-periodic with configurable inter-packet arrival rate] (e.g. 60 to 500Hz)   + Traffic file distribution: [Uniform distribution with configurable packet size] (e.g. 30 to 250B)   + Number of data streams: [Configurable number of streams, configurable traffic parameters common to all streams] (e.g. single/multiple streams with bounded latency)   + Traffic parameters of each data stream: [Configurable data rate, latency and reliability] (e.g. 500kbps, 10ms, 10E-04 PER) |
| Xiaomi | Proposal 1: Periodic traffic can be assumed for the DL and UL traffic of the VR service |
| Apple | Proposal 1: It is key to include uplink traffic with substantial throughputs in the study of AR2.  Proposal 3:   * In the traffic model for XR, multiple data flows (e.g. for audio and video) for each direction (DL or UL) are generated for a UE; * Each data flows can be configured separately with * Periodicity * Packet size distribution (e.g. fixed or following a distribution) * Data flow specific latency and reliability requirements   Proposal 4:   * For XR conversational, for each data flow, the following can be configured separately:   + Periodicity   + Packet size distribution (e.g., fixed or following a distribution)   + Data flow specific latency and reliability requirements * downlink traffic includes the following data flows:   + 2D video   + Audio   + Data stream * uplink traffic includes the following data flows:   + Video + Depth   + Front facing Camera   + 3/6 DOF Pose   + Audio   + Data stream |
| QC | Proposal 3: RAN1 supports the following statistical model for pose/control traffic in UL for XR evaluation.   * Data rate is given. * File is generated periodically. * File size is fixed.   Proposal 4: RAN1 supports the following statistical model for scene update traffic in UL for XR application.   * Data rate is given. * File is generated periodically. * File size is random following truncated Gaussian distribution * File need to be transmitted within file delay budget |
| Nokia, NSB | ***Proposal 2****: For VR1, consider a periodic traffic model as an uplink traffic model. Assume a constant inter-arrival time of 1 packet every 100 ms, where the average bitrate is equal to 8 kbit/s.*  ***Proposal 8:*** *For AR1, consider the FTP Model 3 as an uplink traffic model for Option 1 and Option 2. For Option 1, assume a traffic source generating 58 Mbit/s and 90 Mbit/s as a bitrate for 1080p (Full HD) and 4K video quality, respectively. For Option 2, assume a traffic source generating 10 Mbit/s as a bitrate for 720p video quality.*  ***Proposal 12:*** *For CG, consider the Periodic traffic model as uplink traffic model.* *Assume an inter-arrival time of X/bitrate [seconds] where X is the packet size for uplink transmissions. A transmission bitrate of 2 kbit/s can be assumed.* |

**Summary**

* No UL modeling required: Oppo
* UL modeling required: vivo, InterDigital, Xiaomi, Apple, QC

Given that UL pose/control is common UL traffic for VR/AR/CG, we make following proposal.

**Proposal 11**: RAN1 adopts following UL traffic model for pose information for VR2/AR2/CG

* Periodicity: [2/4/8/16ms]
* Packet size: 100byte
* PDB: 10ms

**Question 11**. Please share your comments on Proposal 11.

|  |  |
| --- | --- |
| Company | View |
| OPPO | We are ok with the evaluation of UL traffic if majority companies support it. |
| MTK | * Periodicity: 4 ms * Packet size: 61 bytes * PDB: 10ms |
| Xiaomi | We prefer to a single periodicity value 4ms. Agree with others. |
| QC | Considering simulation workload, we can evaluate two values of pose information periodicity, 2 and 16.7ms. Packet size of 100bytes and PDB of 10ms seem reasonable. |
| Nokia, NSB | We propose to agree on the number of flows per each direction first and only then proceed with discussion on the possible parameters. |
| CATT | We are OK with the principle of Proposal 11. However, we need to consider the SR and scheduling delay |
| Futurewei | Agree with Nokia that we should first agree on the number of traffic flows on each direction. If the group agrees on single traffic flow modelling on UL, this proposal might not be necessary. Furthermore, our opinion is that evaluation/modelling of DL should take higher priority than UL. |
| InterDigital | We are fine with FL’s proposal on periodicity and PDB. For packet size we prefer using values [100B, 250B] to account for different resolution in pose information for XR/CG applications |
| DOCOMO | We support the proposal in general but prefer single value for the periodicity with 4 ms. |
| ZTE, Sanechips | We prefer to focus on DL for this meeting. |
| AT&T | Support the proposal |
| vivo | Support Proposal 11 in principle. The values for periodicity can be down-selected to reduce evaluation efforts, e.g. 2ms or 4 ms. |
| Huawei, HiSilicon | For a given XR or CG application, different roundtrip interaction delays can result in different user experiences. To reflect different levels of user experience, more PDB values can be evaluated, e.g., 5ms, 20ms.  We suggest to discuss multiple data steam traffic model in Section 2.2.6 first, and then come back to the details of each data stream. |
| Sony | UL should also be as important as DL in this study. Generally, we are in line with Apple view on proposal 1/3/4. |
| LG | Okay in principle with the proposal. Further down selection on the periodicity is preferred. Either 4 or 2 ms is okay to us.  Nokia’s suggestion on the order of discussions is also fine to us. |
| Ericsson | Essentially support, but it is not clear to us if it’s relevant to evaluate capacity for this type of traffic. Highest values of periodicity can be mandatory. |

UL scene (camera feed) is additional UL traffic for AR application.

**Proposal 12**: RAN1 adopts following UL traffic model for UL scene (camera) information for AR

* Bit rate: 10, 20Mbps
* Periodicity: 16.67ms
* PDB: 60ms

**Question 12**. Please share your comments on Proposal 12.

|  |  |
| --- | --- |
| Company | View |
| OPOO | We have the similar comment as for DL traffic that 120fps should be included for better future-proof |
| MTK | Video (scene) traffic (1080x720, 30 fps)   * Interval: 33 ms * Size: 10k bytes (2.4Mbps) * PDB: 40ms |
| QC | We support the modelling UL scene in Proposal 12. |
| Nokia, NSB | We propose to agree on the number of flows per each direction first and only then proceed with discussion on the possible parameters. |
| CATT | We are OK with the proposal. |
| Futurewei | Agree with Nokia that we should first agree on the number of traffic flows on each direction. Furthermore, our opinion is that evaluation/modelling of DL should take higher priority than UL. |
| InterDigital | We are fine with FL’s proposed values |
| DOCOMO | Similar to Question 4, 120 fps can be considered optionally considering CG applications. |
| ZTE, Sanechips | We prefer to focus on DL for this meeting |
| AT&T | 120fps should also be simulated |
| vivo | Suggest to use same traffic model as AR DL Bit rate: 30Mbps, Periodicity: 16.67ms, FFS: PDB. |
| Huawei, HiSilicon | Similar to our view in Question 11, to reflect different levels of user experience, more PDB values can be evaluated, e.g., 30ms, 100ms.  We suggest to discuss multiple data steam traffic model in Section 2.2.6 first, and then come back to the details of each data stream. |
| Sony | SA4 mentioned 2 Cameras for uplink scene 1 camera with bit rate 10, 20 Mbps and 1 camera with 3 Mbps. At least we should discuss whether we want to simulate with both cameras or only one. Proposal 12 above is basically assuming to use 1 camera. |
| LG | Okay in principle with the proposal. FFS exact values of periodicity and bit rate. |
| Ericsson | Support. Lower value can be mandatory. |

## Per UE KPI (whether UE is sa

During RAN1 103-e, the following agreement was made.

|  |
| --- |
| **Agreement 6**: System capacity is defined as the maximum number of users per cell with at least X % of UEs being satisfied.   * X=90 (baseline) or 95 (optional) * Other values of X can also be evaluated optionally   Note: The exact ‘satisfied’ requirements will be discussed separately  FFS: how to calculate the percentage of satisfied users across multiple drops of simulations |

In this section, we discuss how to define the ‘satisfied’ requirements. This completes the system capacity definition, according to the above agreement.

### Capacity KPI

|  |  |
| --- | --- |
| Company | View |
| FutureWei | *Proposal 7: The capacity is defined as the maximum number of users that can be supported with percentage of users that satisfy both latency and PER requirements above a threshold (e.g., 90%)* |
| Oppo | Proposal 9: For each identified traffic/service, whether a UE is satisfied or not is not determined based on the following tuples, where the detailed values is to be determined based on the traffic models.   * DL: {Data rate, Packet Delay Budget, Packet Error Rate} * UL: {Data rate, Packet Delay Budget, Packet Error Rate}   Proposal 10: For each identified traffic/service, the following results are provided   * CDF of achievable data rate * CDF of packet delay |
| Huawei | *Proposal 7: RAN1 needs to identify a KPI that can reflect the user experience in XR and CG services*   * *The identified KPI can reflect the impact of network transmission on* *the user experience.* * *The identified KPI can be calculated with RAN available information.*   *Proposal 8: The identified KPI can be calculated with RAN available information, such as packet loss information, packet delay information, and some XR/CG source related information if they can be available within RAN.* |
| CATT | *Proposal 8: For XR service evaluation, the latency could be defined as the delay budget of Air interface is 20*%*~25% of end-to-end latency requirement.*  *Proposal 9: For XR service evaluation in RAN1, the reliability could be specified by mapping of end-to-end PER of XR service requirements to BLER in Uu interface.*  *Proposal 10: For XR service evaluation, short term throughput, which is be defined as throughput of MAC PDU within delay budget, could be considered as the UPT of XR services.* |
| vivo | *Proposal 3: A UE is regarded as satisfied if the packet error ratio measured for it is equal to or less than the given PER.*  *Proposal 4: The following metrics can be considered for XR capacity evaluation,*   * *Percentage of satisfied UEs* * *System capacity* * *CDF of packet error ratio* * *CDF of packet latency* * *CDF of user-perceived throughput* * *Resource utilization*   *Proposal 5: Percentage of UEs being satisfied for each drop can be calculated separately, and then averaged over all the drops.*  *Proposal 7:* *The user interaction delay can be used as a key metric for uplink capacity evaluation for uplink interaction and pose information traffic.*  *Proposal 8:* *The number of satisfied users for interaction and pose information is defined as the maximum number of users per cell for which the A%-tile user interaction delay is equal to or less than the uplink PDB, where the threshould A% should be discussed and determined, when only interaction and pose information are modelled in uplink.* |
| ZTE | 1. UE is satisfied if the packet loss rate is less than 10-3 or 10-4 for VR2, AR2 and CG.  * Delay threshold for VR2 and AR2 is 60ms and delay threshold for CG is 80ms. * Note: The packet loss rate for both left and right eye should be less than 10-3 or 10-4 if traffic for two eyes is considered |
| InterDigital | Proposal 1: Given the XR applications, define user experience satisfaction by the percentage (70%, 90%) of achieved maximum throughput, which is derived from the traffic parameters of the XR application.  Proposal 2: Evaluate capacity for XR applications by analyzing CDF of UEs achieving an average throughput that is at least equal to the required throughput for the given XR application. |
| MTK | Proposal 8: For XR/CG evaluation, if one frame is mapped to several packets (Ex. constant packet size), then the packet size, and mapping between PDB/PER and frame error rate/QoS needs to be further clarified.  Proposal 9: For XR/CG evaluation, if one frame is mapped to one packet, a UE is satisfied if 99% files are delivered within packet delay budget (PDB) = 10ms. For requirements defined in SA2 LS S2-2009227 (5ms/99.9% and 10ms/99.9%), they can serve as optional (advanced) user satisfaction criterion. |
| Intel | * *Proposal-2: Consider defining the following KPIs for capacity evaluations:*   + *Average data-rate requirement*   + *Packet delay statistics and Packet delay budget (PDB)*   + *Average packet error rate (PER) statistics and reliability requirement*   + *User satisfaction ratio* |
| Nokia | **Proposal 8:** *The percentage of satisfied users from a simulation campaign with N separate drops of M users is calculated for the NxM samples of user satisfaction.*  Proposal 9: *Distinguish UL and DL KPIs and evaluations, allowing conclusions such as “Up to 10 XR devices running VR1 can be supported in a given scenario, while the corresponding UL traffic can only be supported for no more than 7 VR1 XR devices.”*  Proposal 10: *Adopt the following definition for a satisfied UE: “A satisfied UE operates with target link reliability R under L latency bound rate.” The exact values for R, L can be determined after the traffic model and the use case are agreed.*  Proposal 11: *Reliability is defined as a percentage of (DL or UL) packets delivered within a certain packet delay budget (PDB): fraction of (DL or UL) packets that have been correctly received within the preconfigured PDB with respect to the packets generated by the XR application during the simulation time.*  Proposal 12: *Latency bound L is defined as a packet delay budget (PDB), where a packet delay is measured as the difference between the time the packet is received at Layer-2 at the transmitter and the time when it is forwarded from Layer-3 at the receiver end to the upper layers. To facilitate the comparison of results reported by different companies, we propose to set the core network delay to zero.* |
| Xiaomi | Proposal 3: Packet loss rate and delay can be used as the criteria to identify whether UE is satisfied.  Proposal 4: DL capacity and UL capacity can be evaluated separately |
| Samsung | *Observation 3:* *System capacity for a XR application corresponds to a % of UEs for which a target data rate with a latency bound for the XR application is achieved.*  *Observation 4:* *XR can re-use the definitions for system capacity and % of satisfied UEs used in URLLC.*  Proposal 9: PDCCH blocking is part of the XR SI in the evaluation of the latency KPI. |
| Ericsson | Proposal 4: The fraction of satisfied users subject to frame latency bound with a reliability target should be considered as the main system performance measure.  Proposal 5: Latency and reliability metrics for XR use cases should be measured per application PDU but exclude latency and errors contributed by non-RAN aspects such as application, core network, and transport layer.  Proposal 6: RAN1 should decide exact parameters for latency and reliability criteria further when ongoing SA WG4 XR study is finalized [3]. |
| Apple | Proposal 4: propose a UE is deemed satisfied if the QoS requirements for all its data flows are met. |
| QC | Proposal 2: For XR DL evaluation, use following performance metrics for evaluation.   * File error rate and a corresponding threshold (99%) to determine satisfaction of the quality condition * Transmission delay * FFS: Percentage of damaged area and a corresponding threshold to determine satisfaction of the quality condition   Proposal 3: For XR UL scene upload evaluation, use following performance metrics for evaluation.   * File error rate and a corresponding threshold to determine satisfaction of the quality condition * Transmission delay * FFS: Percentage of damaged area and a corresponding threshold to determine satisfaction of the quality condition   Proposal 4: For XR UL pose/control info evaluation, use following metrics for evaluation.   * Age of pose (AOP) |

**Summary**

* Definitions of satisfied UE
  + Satisfying reliability, latency requirement

Based on views from companies we make following proposal.

**Proposal 13.** A UE is declared a **satisfied** UE if more than X=99(%) of packets are successfully transmitted within a given packet delay budget (PDB), where packets which cannot meet its PDB requirement are counted as lost (dropped from Tx queue).

**Question 13**. Please share your comments on Proposal 13.

|  |  |
| --- | --- |
| Company | View |
| OPPO | Support FL’s proposal |
| MTK | Agree on Proposal 13 |
| Xiaomi | Agree with FL proposal. |
| QC | We support proposal 13. This is a clean definition capturing both reliability and latency requirements. In practice, it may be up to gNB implementation whether to server packets with a long delay in queue/buffer. However, in our evaluation, given that we assume a fixed PDB per packet for over the air transmission – this may not be the case in practice, **it may be a reasonable assumption to drop packets from the queue/buffer that have stayed longer than PDB**. In addition, having common assumptions among companies would help to reduce variance of results from different companies. |
| Nokia, NSB | We propose to naturally separate the proposal on UE satisfaction metric from the modelling aspects. We also propose to leave the exact value of X% until the decision on traffic has been made. Therefore, we suggest the following modified proposal:  **Proposal 13.** A UE is declared a **satisfied** UE if more than X (%) of packets are successfully transmitted within a given packet delay budget (PDB). The exact value of X is FFS. |
| CATTT | We are OK with Proposal 13. |
| Futurewei | We support Proposal 13. |
| InterDigital | We think Proposal 13 can be considered as one of the KPIs for evaluating user satisfaction. However, a more comprehensive definition of satisfied UE should also account for satisfying a target data rate requirement along with PDB and PER. |
| DOCOMO | OK with the proposal. |
| ZTE, Sanechips | Support. |
| AT&T | This OK when considering multiple DL/UL flows for a given UE simultaneously and per-flow satisfaction should also be considered. |
| vivo | We support FL’s Proposal 13. X=99% can be baseline, other values can be optional e.g.99.9% or 99.99%. |
| Huawei, HiSilicon | Disagree.  In RAN1#103-e, the following agreement on capacity KPI was achieved:  Agreement:  The following aspects are to be discussed after traffic model is stable.   * For the system capacity definition, how to determine whether a UE is satisfied or not is to be deferred until the exact traffic model along with how to measure E2E user experience is available.  Additional metrics to be collected will be further discussed after traffic model is stable. * Various options for traffic arrival offset among UEs per cell were proposed by companies, e.g., even offset, random offset, no offset. It will be discussed after traffic model is determined.   According to the agreement, how to determine whether a UE is satisfied or not should be considered along with traffic model and E2E user experience, which is not reflected in the current proposal 13.  In fact, the current proposal seems to be a URLLC capacity KPI and does not consider the key characteristic of XR services, e.g., multiple data stream traffic model, user experience, etc.  PER/PDB on its own are not enough to reflect user experience. Given the same PER, different error pattern will result in different user experience, and this is also reflected in the SA4 outcome.  For example, based on the information of P-Trace provided by SA4 (especially the red parts in Table 1 below), the importance of each packet and the mapping relationship between packet and slice or frame can be available in RAN1 from RAN1 evaluation perspective. The packet importance refers to the importance of the slices and the frames which the packet belongs to. Different importance of the slices and the frames have different impact on the user experience. Therefore, in RAN1 performance evaluation, even with the same packet error rate, the error of different packets may result in different user experience since they may belong to different slices or frames which have different importance. Therefore, if some XR/CG source related information, e.g. the mapping between packet and slices or frames and the packet importance, can be available within RAN and is considered in the KPI, the KPI can reflect the impact of network transmission on user experience more accurately.  So in addition to PER/PDB, we think such XR/CG source related information and multiple data streams traffic model should also be considered when determining a XR/CG user is satisfied or not. We suggest to agree the following proposal as a starting point, and RAN1 can continue discussing the details in the remaining time of RAN1#104-e.  *Proposal: The following factors are considered when determining a XR/CG user is satisfied or not:*   * *Packet loss information* * *Packet delay information* * *Some XR/CG source related information if they can be available within RAN, e.g. the mapping between packet and slices or frames and the packet importance* * *Multiple data streams traffic model*   Table 1 P-Trace format table in S4aV200631: P-Trace format   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | *(copied from S4aV200631 (attachment in SA4’s LS to RAN1))*  … 3.3.4 P-Trace For each packet in the delivery, the following information is provided.   |  |  |  | | --- | --- | --- | | Name | Type | Semantics | | number | BIGINT | Unique packet number in the delivery | | time\_stamp\_in\_micro\_s | BIGINT | Availability time of packet for next processing step relative to start time 0 in microseconds (0 means lost). | | size | BIGINT | packet size in bytes. | | user\_id | BIGINT | assigns an id to the user in order to differentiate | | buffer | BIGINT | The associated eye buffer 1=left 2=right  In general, differentiates application traffic for different buffers, for example audio, video, left eye, right eye. For example mapped to port or track. | | delay | BIGINT | Delay observed of the packet in the last processing step (-1 means lost) | | render\_timing | BIGINT | the rendering generation timing associated to the media included in the packet. | | number\_in\_unit | BIGINT | The number of the packet within the unit (slice), start at 1 | | last\_in\_unit | BIGINT | Indicates if this is the last packet in the slice/unit 0=no, 1=yes | | type | BIGINT | The data type of the unit  0 unknown  For video 1=intra 2=inter | | importance | BIGINT | assigned relative importance information (higher number means higher importance) | | index | BIGINT | Unique index increased by 1 and indexing this row in the S-Trace file. | | s\_trace | STRING | Reference to s\_trace file containing information for each slice |   … | |
| Sony | Support Proposal 13 |
| LG | Okay with the proposal. |
| Ericsson | Is there a typo? Packets should be successfully received, not transmitted. Also clarify that by packet does not mean “IP packet”. Otherwise OK. |

# References

1. R1-2101765 LS to on XR-Traffic Models, SA4

18.4.1

1. [R1-2100055](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100055.zip) XR traffic model, FUTUREWEI
2. [R1-2100132](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100132.zip) Discussion on the XR traffic models for evaluation, OPPO
3. [R1-2100207](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100207.zip) Discussion on applications and traffic model for XR and Cloud Gaming, Huawei, HiSilicon
4. [R1-2100361](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100361.zip) XR traffic model, CATT
5. [R1-2100476](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100476.zip) Discussion on traffic models of XR, vivo
6. [R1-2100528](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100528.zip) Discussion on Traffic Model for XR evaluations, ZTE , Sanechips
7. [R1-2100555](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100555.zip) Discussion on traffic model for XR study, LG Electronics
8. [R1-2100571](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100571.zip) Discussion on XR applications and traffic models, InterDigital, Inc.
9. [R1-2100680](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100680.zip) On traffic model for XR, Intel Corporation
10. [R1-2100724](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100724.zip) On Traffic Model for XR study, Nokia, Nokia Shanghai Bell
11. [R1-2100775](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100775.zip) XR Traffic Model Considerations, AT&T
12. [R1-2100879](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100879.zip) Discussion on XR Applications and Evaluation Assumptions, Sony
13. [R1-2101101](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101101.zip) Discussion on Traffic model for XR evaluation, Xiaomi
14. [R1-2101137](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101137.zip) Traffic Model for XR and CG, MediaTek Inc.
15. [R1-2101240](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101240.zip) XR Applications and Traffic Models, Samsung
16. [R1-2101314](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101314.zip) Traffic model for XR, Ericsson
17. [R1-2101365](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101365.zip) Views on XR traffic models, Apple
18. [R1-2101493](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101493.zip) XR Traffic Models, Qualcomm Incorporated
19. [R1-2101635](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101635.zip) Discussion on traffic model for XR, NTT DOCOMO, INC

8.14.2

1. [R1-2100056](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100056.zip) XR evaluation methodology FUTUREWEI
2. [R1-2100133](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100133.zip) Discussion on the XR evaluation methodology OPPO
3. [R1-2100242](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100242.zip) Discussion on evaluation methodology for XR and Cloud Gaming Huawei, HiSilicon
4. [R1-2100362](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100362.zip) Evaluation methodology and performance index for XR CATT
5. [R1-2100477](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100477.zip) Discussion on evaluation methodologies of XR vivo
6. [R1-2100529](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100529.zip) On XR Evaluation Methodology ZTE , Sanechips
7. [R1-2100556](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100556.zip) Discussion on evaluation assumption for XR study LG Electronics
8. [R1-2100572](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100572.zip) Discussion on Evaluation Methodology for XR InterDigital, Inc.
9. [R1-2100586](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100586.zip) On Evaluation Methodology for XR and CG MediaTek Inc.
10. [R1-2100681](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100681.zip) On evaluation methodology for XR Intel Corporation
11. [R1-2100725](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100725.zip) Development of the Evaluation Methodology for XR Study Nokia, Nokia Shanghai Bell
12. [R1-2100776](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2100776.zip) XR Evaluation Assumptions AT&T
13. [R1-2101102](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101102.zip) Discussion on evaluation methodology for XR services Xiaomi
14. [R1-2101241](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101241.zip) XR Evaluation Methodology and KPIs Samsung
15. [R1-2101315](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101315.zip) Evaluation methodology for XR Ericsson
16. [R1-2101366](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101366.zip) Views on XR evaluation methodology Apple
17. [R1-2101494](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101494.zip) Evaluation Methodology for XR Qualcomm Incorporated
18. [R1-2101636](https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_104-e/Docs/R1-2101636.zip) Discussion on evaluation methodology for XR NTT DOCOMO, INC.