3GPP TSG RAN WG1 #105-e R1-21xxxxx

e-Meeting, May 10th – 27th, 2021

Source: Moderator (Qualcomm)

Title: [104b-e-NR-XR-01] Email discussion/approval on traffic model

Agenda Item: 8.14.1

Document for: Discussion and Decision

# Introduction

This contribution discusses traffic models based on the tdocs that have been submitted to RAN1#105e [1-18].

# PDB for UL AR

Below are RAN1 agreements on UL AR traffic model. As noted, PDB values in [ ] for Stream 2 in Option 1 and 3, and Option 2 are to be further discussed and potentially confirmed in RAN1#105-e, where other values can be also discussed if needed.

Agreements:

For evaluations of AR in UL:

* Option 1 (Baseline for power and capacity evaluations): Two streams as defined below
  + Stream 1: pose/control
    - Traffic model and QoS parameters are same as for pose/control for UL CG/VR.
  + Stream 2: A stream aggregating streams of scene, video, data, and audio.
    - Packet size: Truncated Gaussian distribution with the parameter values same as for DL
    - Periodicity: 60 fps
      * Jitter (optional): same model as for DL
    - Data rate: 10 Mbps (baseline), 20 Mbps (optional)
    - PDB: [60] ms (baseline), [10/15] ms (optional)
* Option 2 (Optional for power evaluation and baseline for capacity evaluation): Single stream as defined below
  + Packet size: Truncated Gaussian distribution with the parameter values same as for DL
  + Periodicity: 60 fps
    - Jitter (optional): same model as for DL
  + Data rate: 10 Mbps (baseline), 20 Mbps (optional)
  + PDB: [60] ms (baseline), [10/15] ms (optional)
* Option 3 (Optional): Three streams as defined below
  + Stream 1: pose/control
    - Traffic model and QoS parameters are same as for pose/control for UL CG/VR.
  + Stream 2: A stream aggregating streams of scene and video
    - Packet size: Truncated Gaussian distribution with the parameter values same as for DL
    - Periodicity: 60 fps
      * Jitter (optional): same model as for DL
    - Data rate: 10 Mbps (baseline), 20 Mbps (optional)
    - PDB: [60] ms (baseline), [10/15] ms (optional)
  + Stream 3: A stream aggregating streams of audio and data
    - Periodicity: 10ms
    - Data rate: 0.756 Mbps/s or 1.12 Mbps
    - Packet size: determined by periodicity and data rate
    - PDB: 30 ms
* Option 4 (Optional): Three streams as defined below
  + Stream 1: pose/control
    - Traffic model and QoS parameters are same as for pose/control for UL CG/VR.
  + Stream 2: I-stream for video
  + Stream 3: P-stream for video
  + Note: For stream 2 and stream 3, the I/P-stream model for DL video can be reused for UL video.  Companies should report detailed assumptions in their simulations on packet size distribution for each stream, packet arrival interval (or fps) for each stream, PDB for each stream, PER requirement for each stream, criteria to be satisfied UE.
  + Companies should strive to align the parameter values for the options chosen as much as possible
* Note: Above PDB values in [ ] for Stream 2 in Option 1 and 3, and Option 2 are to be further discussed and potentially confirmed in RAN1#105-e, where other values can be also discussed if needed.
* In case multiple steams are evaluated for UL AR, a UE is declared as satisfied only when each stream meets the requirement that X (%) of packets are successfully delivered within a given air interface PDB.
  + X value for pose/control: follow X values for pose/control for CG/VR
  + X value for other stream: follow X values for DL video stream.

SA4 studies on characters and requirements for AR2 are summarized in the following table.

Characters and requirements for AR2 in UL (S4aV200640)

|  |  |  |
| --- | --- | --- |
| **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 |

**Companies’ views in tdocs [1-18] on PDB of the stream aggregating streams of scene, video, data, and audio**

* 60 ms (6 companies): FUTUREWEI, QCOM, Intel, Samsung, ZTE, Ericsson
* 10 ms (5 companies): CATT, Apple, LG, InterDigital, DCM
* 15 ms (4 companies): CATT, OPPO, LG, DCM

**Question 1. Please share your view on the PDB value of the stream in UL AR aggregating streams of scene, video, data, and audio, i.e., Stream 2 in Option 1 and 3, and Option 2. FL is planning to propose majority view as baseline.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
| LG | Referring to the modelling assumption in SA4 (S4-210614), the maximum latency for slice for AR2 UL is 60ms in 8.4.5, and 80ms in 9.2.1. As the maximum latency for slice in downlink is 60ms for VR2/AR2, and 80ms for CG, we think assuming the air PDB value for AR2 UL to be the same as that for either VR2/AR2 or CG should be a natural consequence. Therefore, we support 10/15 ms as baseline air PDB values, and 60 ms as optional. |
| Apple | Our preference is 10 ms, and LG’s observation is valid. As a compromise we are also fine with 15 ms. Hope other companies supporting 60 ms would be fine with 10 ms or 15 ms. |
| QC | The two AR UL flows: 1) pose and 2) UL camera/video/voice should have clearly different latency requirement for the following reasons.   * 3/6DOF Pose captures the users (head) motion and direction of view, which is sent to server and used to render a new scene which is in line with users view port. If the pose is sent with delay, what the user see in his/her display would be lagging compared to his/her movement and the user will see misalignment between rendered frame and his movement/ or real environment through the glass. This could reduce impressiveness or can cause dizziness. Thus, the feedback loop from pose to render to display (which we typically call motion to render to photon(eye) M2R2P) should be quite tight and normally in the range of 60~70ms. The UL transfer delay budget in air-interface is assumed to take 10ms (according to our agreement). * The other UL traffic (including camera, data, voice etc.) are information for conversational purpose. The camera/video are what are captured by front facing camera installed in the users HMD/AR glasses. It is environment and/or users movement and sent to other users through central server. The nature of this conversational traffic is different from 3/6DOF Pose. In conversational XR session between two users A and B, user A’s camera/video information does **not depend on** user B’s motion. Therefore, user A’s camera/video/voice does not need to be sent as fast as pose info in user A’s device. In user A’s display, a delayed user B’s image could still be rendered. As long as the camera/video/voice are sent within the conversational latency requirement, both users will not feel quality degradation. Due to this different nature, the e2e latency requirement of such traffic is higher than M2R2P. As captured by SA4, in S4aV200640, it is in the range of 100ms, 200ms. As a reference, the mouth-to-ear delay to support interactive service is around 150ms.   **Based on this reason, we think the traffic for interactive conversation could have larger latency requirement of 60ms than that of Pose.**  Note that TR26.928 does not have specific values for XR conversational traffic’s UL PDB. They are still FFS. Check Table 6.3-1. |
| Samsung | We have the same observations as QCOM. (UL) Pose has much more stringent PDB requirements than UL video, voice or camera. Pose “motion-to-photon” latency must be in the order of 60 ms or it becomes perceptible (and unpleasant) to the user. Therefore, the 10ms UL transfer budget is appropriate. UL voice, video or camera are real-time conversational in nature and supplement pose processing on the network side, they’re not key to rendering for the DL. Therefore, the traditional 200ms end-to-end delay is more applicable with 60ms allocated to the UL transfer budget. |
| InterDigital | We have a similar understanding with LG and Apple regarding the air-interface latency for the AR aggregated stream. The prefer the PDB value to be either 10ms or 15ms as baseline. |
| vivo | We have the similar view as QC. For UL video stream for AR application, it is captured by sensors such as a micro camera in the AR device, and then rendered, compressed and transmitted to the server side. Compared with the pose/control stream, the E2E latency requirements of video or scene information are much looser as shown in the table provided in S4aV200640. Assuming 100-200 ms E2E delay, then the air-interface PDB requirement for UL video stream could be relaxed to 60ms. |
| OPPO | According to the SA4 LS, the E2E latency is as blow:   |  |  | | --- | --- | |  | Maximum latency for slice | | VR DL video stream | 60ms | | CG DL video stream | 80ms | | AR DL video stream | 60ms | | AR UL video stream | 80ms |   As we know, the air-interface latency is only a portion of the whole E2E latency. As a result, RAN1 agreed an air-interface PDB of 10ms and 15ms in RAN1 evaluation corresponding to the E2E latency of 60ms and 80ms for DL video stream. Following the similar correspondence between air-interface PDB and E2E latency, the value of 15ms should be used for the air-interface PDB of AR UL video stream. By the way, we can accept 10 ms as well. |
| CATT | The PDB in RAN1 should be only fraction of end-to-end PDB, which is around 20-25% based on 3GPP evaluation assumption. Thus, the XR PDB in RAN1 should be around 10-15 ms. |
| DOCOMO | We share similar understanding with LG and Apple but at the same time, we think the observation of QC is also valid. We are OK with 60 ms as a compromise. |
| MTK | We share similar view with QC/vivo and think 60ms PDB can be the baseline for AR UL video. |
| Ericsson | We understand the argument that the latency requirement for pose is more stringent than the latency requirement for the aggregated AR flow. Then we also acknowledge the arguments from LG and Apple that 60ms would seem to be somewhat high.  Could we aim for something like 40ms as baseline and keep 10/15ms as optional? If not, we prefer 60ms. |
| Xiaomi | We share the similar understanding as QC that the UL video/audio traffic for AR2 are mainly for conversation purpose, as shown in table of requirements for AR2 in UL. |
| ZTE, Sanechips | We agree with Qualcomm’s opinion. The PDB of aggregating streams (video / scene) or the single video stream should be set differently with that of Pose/Control stream, due to their different characteristics. According to SA4 input, S4aV200640 and analysis in R1-2105603, the E2E latency requirement of video is 100ms or 200ms. Considering pre-coding delay and encoding delay, 60ms may be a reasonable PDB as baseline for AR UL. |
| Huawei, HiSilicon | In S4aV200640, as the table shows, the AR UL E2E latency could be 60ms, 100ms, 200ms. And as OPPO further pointed out, the AR UL E2E latency could also be 80ms as per SA4 LS.  So probably this value could be very different depending on application configurations, codec, etc.  Since there is a long gap between May meeting and Aug meeting, maybe RAN1 does not need to decide one value as baseline. RAN1 can encourage companies to simulate both a long PDB (60ms) and a short PDB (10 or 15ms), such results are informative since RAN1 can know how UL PDB can impact the capacity and UE power consumption. |
| Nokia, NSB | We also consider that 60 ms is a bit large for AR UL PDB. Following the same calculation as for VR DL PDB, we can assume that AR UL has a PDB equal 10 or 15 ms. |

# Dual Eye Buffer for DL video stream

As indicated below, it is FFS whether and how to evaluate single eye and dual eye buffer.

|  |
| --- |
| RAN1 Agreement:  Parameters of Truncated Gaussian distribution for packet size of DL video stream in case of single stream evaluation (note: these parameter values are those before the truncation):   * [STD, Max, Min]: [10.5, 150, 50]% of Mean packet size * Other values that can be used for evaluation: [STD, Max, Min] = [4, 112, 88] % of Mean for single eye buffer, [3, 109, 91] % of Mean for dual eye buffer * FFS: Whether and how to evaluate single eye and dual eye buffer * Note: Companies report the values used in their simulation results. * Note: There is no consensus that the [10.5, 150, 50]% of mean packet size is the best set of parameters |

Company views on the FFS point are summarized in the table.

|  |  |
| --- | --- |
| **Company** | **Proposals in tdocs** |
| vivo [3] | *Proposal 1: For dual-eye buffer, the dual-eye buffer traffic model in Table 1 can be optionally evaluated.*  Table 1. Dual-eye buffer traffic model   |  |  |  |  | | --- | --- | --- | --- | | **Traffic model** | Single stream | Dual-eye buffer | Note | | **Date rate (Mbps)** | 45 | 45 |  | | **Packet size distribution** | Truncated Gaussian distribution | |  | | **Mean packet size (Bytes)** | 93750 | 46875 | Average data rate / FPS / 8 [bytes] | | **STD of packet sizes (Bytes)** | 9844 | 4922 | 10.5% \* Mean packet size | | **Maximum packet size (Bytes)** | 140625 | 70312 | 150% \* Mean packet size | | **Minimum packet size (Bytes)** | 46875 | 23437 | 50% \* Mean packet size | | **Packet arrival interval (ms)** |  |  |  | | **Packet delay budget (ms)** | 10 | |  | |
| CATT [4] | ***When single eye or dual eye buffer is modelled for evaluation, the total transmission data rate would be kept the same with the details of modelling reported by companies.*** |
| Samsung [11] | *The dual-eye buffer case is evaluated as single stream case by doubling the FPS and halving the packet size compared to the single-eye buffer case.* |
| LG [13] | *Proposal 1: Single and dual eye buffer can be modelled using the baseline single-stream DL traffic model with the agreed baseline and optional parameters.*   * *No further discussion on details is needed.* |
| Xiaomi [15] | **Proposal 1: For XR DL evaluation, interleaved eye buffer model can be optionally considered.** |
| ZTE [16] | [*Observation 1:* There is maximum 7% capacity gain for single eye traffic in high system load , compared to dual eye traffic in indoor house scenario, when bit rate is 45Mbps.](#_Toc15034)  [*Proposal 1:* The simulation results based on traffic for dual eye buffer is enough to present the system capacity. And companies can provide the results based on traffic for single eye buffer optionally.](#_Toc9207) |

**Question 2. Based on the company inputs, the following is proposed. Please share your view in the below table.**

**FL proposal: For DL video stream, the dual-eye buffer case can be optionally evaluated, based on the single stream model by doubling the packet arrival rate and halving the packet size compared to sin-eye buffer, while all other parameters (e.g., jitter, PDB) are the same as for single stream. For companies who are evaluating the performance of dual-eye buffer in addition to single-eye buffer (baseline), the following scenarios in the table are recommended as a common baseline (so as to have more results in the same scenario). It is up to company to evaluate other scenarios.**



|  |  |
| --- | --- |
| **Company** | **Comment** |
| LG | Okay with the FL proposal. |
| QC | Support FL proposal. |
| Samsung | We support the FL proposal. |
| InterDigital | We support FL’s proposal |
| vivo | Support FL’s proposal. |
| OPPO | “**doubling the packet arrival rate**” seems not accurate. In fact, the packet arrival rate should be the same as that of single-eye buffer. The difference is that simulator should generate two packets each time for dual-eye buffer.  The following is copied from SA4 LS for reference:  ***Observation 1: For XR and Cloud Gaming, 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.*** |
| CATT | We are OK with FL’s proposal |
| DOCOMO | Support FL proposal. |
| MTK | We are fine with FL’s proposal, while this seems to select “Traffic source type 2” quoted by OPPO as baseline for the optional dual-eye buffer evaluation. |
| Ericsson | Support |
| Xiaomi | We support FL proposal. |
| ZTE, Sanechips | Fine with FL proposal. But one thing we have noticed is that the understanding on single/dual eye buffer. From our understanding, “Single eye buffer” and “Dual eye buffer” represent:   * Single eye buffer: The packet of left eye and right eye arrive in turn. * Dual eye buffer: The packets of both eyes arrive at the same time for each frame.   Considering the agreed traffic of evaluations for single stream is video with two eyes information, we suggest replacing the description, “single eye buffer” and “dual eye buffer” , with “In turn arrival” and “Same time arrival” or add some clarification on single/dual eye buffer. Moreover, minimum packet sizes in table are corrected. As a result, the modified table is shown as follows. As to the deployment scenario, is it correct understanding that these two are picked as examples?   |  |  |  |  | | --- | --- | --- | --- | | **Application** | AR/VR 30Mbps, | |  | | **Deployment Scenario** | Dense Urban for FR1,  InH for FR2 | |  | | **Traffic model** | ~~Single stream~~  Same time arrival | ~~Dual-eye buffer~~  In turn arrival | Note | | **Date rate (Mbps)** | 30 | 30 |  | | **Packet size distribution** | Truncated Gaussian distribution | |  | | **Mean packet size (Bytes)** | 62500 | 31250 | Average data rate / FPS / 8 [bytes] | | **STD of packet sizes (Bytes)** | 6563 | 3281 | 10.5% \* Mean packet size | | **Maximum packet size (Bytes)** | 93750 | 46875 | 150% \* Mean packet size | | **Minimum packet size (Bytes)** | ~~46875~~  31250 | ~~23437~~  15625 | 50% \* Mean packet size | | **Packet arrival interval (ms)** |  |  |  | | **Packet delay budget (ms)** | 10 | |  | |
| Huawei, HiSilicon | We are generally fine with FL’s proposal. Some comments as below:  In single stream model, it seems RAN1 will not prioritize any scenario and just leave it to company report. Following the same principle, we suggest to remove the row “Deployment scenario”.  We also suggest to remove the last row “packet delay budget”, since the agreements for single-stream should apply here, i.e. 10ms is the baseline, and companies can also optionally evaluate 7ms, 13ms, etc. |
| Nokia, NSB | We support FL proposal.  In response to **OPPO and ZTE comment**, from the observation OPPO copied (LS SA4 (S4-210614)), there are two different ways to model two-eye buffer traffic. The one proposed by FL is the **Traffic source type 2**: every 1/(2\*X) s - > doubling the packet arrival rate. While the **Traffic source type 1** assumes doubling the packet size. |

# DL: Two Stream Traffic Models: Option 1

Below is RAN1 agreement related to evaluation of two streams for DL.

|  |
| --- |
| RAN1 Agreement:  In addition to single stream per UE in DL which is baseline, two streams can be optionally evaluated for DL   * Option 1: I-frame + P-frame   + Option 1A: slice-based traffic model   + Option 1B: Group-Of-Picture (GOP) based traffic model * Option 2: video + audio/data * Option 3: FOV + omnidirectional stream * Companies should report detailed assumptions in their simulations on packet size distribution for each stream, packet arrival interval (or fps) for each stream, PDB for each stream, PER requirement for each stream, criteria for being satisfied. * Companies should strive to align the parameter values for the options chosen as much as possible * FFS: Whether audio stream is separate or aggregated with the data stream in option 2 (Intention of option 2 is not to create a 3 stream option) |

Company inputs on the Option 1 are summarized below.

|  |  |
| --- | --- |
| **Company** | **Proposals in tdocs** |
| Huawei [2] | ***Proposal 1: For video of AR/VR/CG, adopt the traffic model in following Table 5 for “Option 1: I-frame + P-frame”.***  Table 5. Option 1: I-frame + P-frame model for DL video   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Two data streams, i.e. M1 = 2** | **Option 1A: slice-based** | | | **Option 1B: GOP-based** | | | | I-stream | | P-stream | I-stream | | P-stream | | **Packet modelling** | Slice-level | | | Frame-level | | | | **Traffic pattern** | Both streams are periodic with the same FPS. | | | Follow the GOP structure, e.g. GOP size K = 8. | | | | **Number of packets per stream at a time** | 1 | | N-1 | 1 or 0 | | 0 or 1 | | * N is the number of slice per frame, e.g. N = 8. | | | | **Average data rate per stream** |  | |  |  | |  | | * R: average data rate of a single stream video * : average size ratio between one I-frame/slice and one P-frame/slice, e.g. = 1.5, 2   + Other values can be optionally evaluated. | | | | | | | **Packet size distribution** | truncated Gaussian distribution | | | | | | | Mean = | Mean = | | Mean = | Mean = | | | * [STD, Max, Min]: [10.5, 150, 50]% of Mean packet size * FPS is the frame rate of the single stream video | | | | | |   Note: the QoS requirement for each stream is separately discussed in the KPI part. |
| vivo [3] | *Proposal 2: Which traffic model to be chosen as the baseline of multi-stream evaluation needs to be further studied.*  *Proposal 3: For multi-stream modelling with I-frame and P-frame, the following traffic characteristics need to be further studied,*   * *The packet size ratio of I-frame and P-frame.* * *The length of GOP for GOP-based model and the number of slices per frame for Slice-based model.* * *The jitter distribution.*   *Proposal 4: How to set the X value and air interface PDB requirements for multi-stream model need to be further studied.*  ***Proposal 5: A UE with multi-stream is declared as a satisfied UE if each stream from the multi-stream has been satisfied, i.e. for each stream more than X (%) of packets are successfully transmitted within a given air interface PDB, where the X value and the given air interface PDB can be set per stream.*** |
| CATT [4] | ***Proposal 2: I-frame and P-frame would not be modelled differently as the baseline.***  ***Proposal 3: The audio stream could be aggregated with the data stream in option 2 (video + audio/data) for two streams per UE in DL.*** |
| Apple [9] | **Proposal 1:**  For DL traffic model Option 2, the audio/data flow is modeled with:   * A stream aggregating streams of audio and data   + Periodicity: 10ms   + Data rate: 0.756 Mbps/s or 1.12 Mbps   + Packet size: determined by periodicity and data rate   + PDB: 30 ms |
| Sony [10] | **Proposal 3: Consider the entire video stream (I-frames, P-frames etc) to be transported on a bearer with a single associated QoS class.** |
| MediaTek  [12] | ***Proposal 1: For the optionally evaluated two-stream DL traffic, prioritize the IDR Group-Of-Picture (GOP) based traffic model, Option 1B) evaluation.***  ***Proposal 2: For the optionally evaluated two-stream DL traffic agreed in RAN1 #104-bis-e, further agree the baseline evaluation parameters about traffic model for Option 1 as in Table I.***  **Table I**   |  |  |  | | --- | --- | --- | | Application | AR/VR/CG | | | Two data streams, i.e. M1 = 2 | * Stream #1: I-stream * Stream #2: P-stream | | | Option 1: slice-based | Option 2: frame-based | | Packet modelling | Slice-level | Frame-level | | Packet size distribution | Truncated Gaussian distribution for each stream. The STD/Max/Min of packet size follows the single stream assumption: [STD, Max, Min]: [10.5, 150, 50]% of Mean packet size | | | Mean packet size | Derived from FPS and average data rate listed below | | | Traffic arrival pattern | Both streams are periodic with 60 FPS. | Follow the GOP structure with GOP size K = 8, i.e., one I-frame is followed by seven P-frames. The overall FPS of I-frame plus P-frame is 60. | | Number of packets per stream at a time | * Stream #1: 1 * Stream #2: N-1 * N is the number of slice per frame, use N = 8. | * Stream #1: 1 or 0 * Stream #2: 0 or 1 * At each time instant, there is either only one I-stream packet or only one P-stream packet | | Average data rate | Stream #1: Stream #2 = : (N-1)  The aggregated data rate of I-stream plus P-stream should be the same as single stream assumptions | Stream #1: Stream #2 = : (K-1)  The aggregated data rate of I-stream plus P-stream should be the same as single stream assumptions | | is average size ratio between one I-frame/slice and one P-frame/slice, use = 3   * Other values can be optionally evaluated. | |   ***Proposal 3: Evaluate [PER\_I, PER\_P] = [1%, 1%] and [0.1%, 5%].***  ***Proposal 4: Evaluate [PDB\_I, PDB\_P] = [10ms, 10ms] and [17ms, 9ms]. It is noted that for GOP size K=8 in Option 1B, [PDB\_I, PDB\_P] = [10ms, 10ms] and [17ms, 9ms] provide the same average PDB.***  ***Proposal 5: RAN1 to coordinate and cooperate with SA4 to construct a video quality evaluation block (as shown in the red block in Figure 3 below) based on statistical models used in RAN1 to evaluate the different QoS requirements and the performance enhancement for various RAN1 proposals.***  ***Proposal 6: Discuss two possible options:***   1. ***FoV vs. non-FoV*** 2. ***FoV vs. low resolution Omnidirectional stream***  * ***FFS: Need for different QoS requirements for the two streams.*** * ***FFS: co-existence with the QoS requirements for I/P-frames.*** |
| LG [13] | ***Proposal 4: For optional two-stream DL traffic models, audio stream is aggregated with the data stream in Option 2*** |
| Xiaomi [15] | **Proposal 2: Send LS to SA4 to ask (X, PDB) requirement and statistical models for packets associated to I-frames and P-frames** |
| ZTE [16] | [*Proposal 3:* Further discussion in RAN1 the parameters of I/P stream modelling for DL video stream and parameters in table 7 can be regarded as starting point.](#_Toc26894)  ***Table 7 Summary of parameters for I/P stream modelling***   |  |  |  | | --- | --- | --- | | ***Application*** | ***AR/VR/CG*** | | | ***Two stream data*** | ***Stream #1: I-frame***  ***Stream #2: P-frame*** | | | ***Option 1 Sliced-based*** | ***Option 2: Frame-based (GoP)*** | | ***Structure*** | ***A frame consists of:***  ***Number of Stream #1: 1***  ***Number of Stream #2: 7*** | ***A GoP consists of:***  ***Number of Stream #1: 1***  ***Number of Stream #2: 7*** | | ***Frame per second*** | ***Stream #1: 60FPS***  ***Stream #2: 60FPS*** | ***Stream #1 + Stream #2 = 60FPS*** | | ***Average packet size ratio*** | ***Stream #1 : Stream #2 = 2:1*** | ***Stream #1 : Stream #2 = 2:1*** | | ***(PSR, PDB)*** | ***AR/VR:***  ***Stream #1: (99%, 20ms)***  ***Stream #2: (90%, 20ms)***  ***CG: Stream #1: (99%, 20ms)***  ***Stream #2: (90%, 20ms)*** | ***Option 1:***  ***Stream #1: (99%, 10ms)***  ***Stream #2: (90%, 10ms)***  ***Option 2:***  ***Stream #1: (99%, 15ms)***  ***Stream #2: (99%, 10ms)*** |   [*Proposal 4:* Further discussion in RAN1 the parameters of FoV and non-FoV stream modelling for DL 360°video stream and parameters in table 8 can be regarded as starting point.](#_Toc31285)  ***Table 8 Initial Parameters of FoV and non-FoV stream modelling***   |  |  |  | | --- | --- | --- | | ***Application*** | ***VR1*** | | | ***Two Stream Data*** | ***Stream #1: FoV stream***  ***Stream #2: Non-FoV stream*** | | | ***Option 1: sliced based traffic model*** | ***Option 2: Two separate streams*** | | ***Structure*** | ***A frame consists of:***  ***Stream #1: 1 (18 tiles)***  ***Stream #2: 1*** | ***A Group of Tiles consist of: Stream #1: 18 tiles***  ***Stream #2: 1*** | | ***Frame Per Second*** | ***Stream #1: 30FPS***  ***Stream #2: 30FPS*** | ***Stream #1: 540 tiles per second***  ***Stream #2: 30FPS*** | | ***Data Rate*** | ***Stream #1: 12.78 Mbps***  ***Stream #2: 8Mbps*** | ***Stream #1: 12.78Mbps (the aggregated data rate of the 18 tiles within a group of tiles)***  ***Stream #2: 8Mbps*** | | ***(PSR, PDB)*** | ***Stream #1: (99%, 20ms)***  ***Stream #2: (90%, 20ms)*** | ***Stream #1: (99%, 10ms)***  ***Stream #2: (90%, 10ms)*** | |
| DOCOMO [17] | **Proposal 2:**   * *Audio stream is aggregated with data stream in option 2 for modeling of DL multiple streams.* |

**FL proposal: Based on the company inputs, the following is proposed.**

**For the optional two stream evaluation of I-frame and P-frame for DL video stream, parameters in the following table are used. For companies who are evaluating this option, the following is recommended (so as to have more results in the same scenario)**

* **Common baseline: AR/VR, 30Mbps (aggregated data rate), Dense Urban for FR1 and InH for FR2.**
* **Companies are also encouraged to evaluate at least, other baseline scenarios/configurations/parameters.**
* **In addition, evaluation of optional scenarios/configurations/parameters is up to company.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Two data streams, i.e. M1 = 2** | **Option 1A: slice-based** | | | | **Option 1B: GOP-based** | | |
| I-stream | | P-stream | | I-stream | | P-stream |
| **Packet modelling** | Slice-level | | | | Frame-level | | |
| **Traffic pattern** | Both streams are periodic at 60 fps with the same jitter model as for single stream. | | | | Follow the GOP structure, where GOP size K = 8 with the same jitter model as for single stream. | | |
| **Number of packets per stream at a time** | 1 | | N-1 | | 1. frame: 1 or 0   P-frame: 0 or 1  At each time instant, there is either only one I-stream packet or only one P-stream packet | | |
| * N = 8 is the number of slices per frame. | | | |
| **Average data rate per stream** |  | |  | |  | |  |
| * R: average data rate of a single stream video * : average size ratio between one I-frame/slice and one P-frame/slice, e.g. = 1.5, 2, 3 | | | | | | |
| **Packet size distribution** | truncated Gaussian distribution | | | | | | |
| Mean = | Mean = | | | Mean = | Mean = | |
| * [STD, Max, Min]: [10.5, 150, 50]% of Mean packet size * FPS is the frame rate of the single stream video | | | | | | |
| **PER, PDB** | [PER\_I, PER\_P] = [A %, B %]  [PDB\_I, PDB\_P] = [C ms, D ms] | | | [PER\_I, PER\_P] = [E %, F %]  [PDB\_I, PDB\_P] = [G ms, H ms] | | | |

**Question 3. Please share your view on the above FL proposal. Please propose values of, A, B, C, D, E, F, G, H (if you are interested in multiple values for a parameter, e.g.** = 1.5, 2, **please provide a single baseline value that can be simulated by more companies, e.g.,**  = 1.5 (baseline), 2 (optional)**.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
| LG | Okay with the FL proposal. |
| QC | We are ok with the FL suggested framework for the evaluation of Option 1. But, regarding the choice of parameters, we suggest discussing next meeting after further study. |
| Samsung | We are in principle fine with the proposal, but think more time is needed to settle on the parameters values. We propose to leave these FFS until August RAN1. |
| InterDigital | We support FL’s proposal |
| vivo | We are fine with FL’s proposal. It is good to have the common framework to proceed the multi-streams evaluation. For the detailed parameters of multi-stream modelling, it seems more discussions are needed. So we are also OK to leave these to the Aug. meeting |
| CATT | We don’t think two stream model of I- and P-frame is necessary since gNB scheduler might NOT have the I-frame and P-frame information. However, we are OK with FL’s proposal if some companies would evaluate it. |
| DOCOMO | OK with the FL proposal. |
| MTK | We support the FL proposal. |
| Ericsson | We do not see the need of a two-stream model. Agreeing on more optional simulation cases does not seem useful in light of the diverging simulation results. |
| ZTE, Sanechips | Firstly, we are fine with the FL proposal.  Secondly, there is a typo in the Option1B of forth row “**Number of packets per stream at a time**” .  ~~P-frame~~ I-frame: 1 or 0  P-frame: 0 or 1  Thirdly, according to the analysis in R1-2105603, the average packet ratio is 2 (baseline) or 1.5 (optional). And the values from A to H are list as follows.   |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** | | 1 | 10 | 20 | 20 | 1 | 1 or 10 | 10 or 15 | 10 |   As a result, the KPI combinations of (PER, PDB) are shown as follows.   |  |  |  | | --- | --- | --- | | ***Application*** | AR/VR 30Mps | | | ***Multi-stream model type*** | Option 1A: Slice-based | Option 1B: Frame-based(GoP) | | ***PER, PDB*** | 1. slice : (1%, 20ms)   P-slice: (10%, 20ms) | Option 1:   1. frame :(1%, 10ms)   P-frame :(10%, 10ms)  Option 2:   1. frame:(1%, 15ms)   P-frame:(1%, 10ms) | |
| Huawei, HiSilicon | We are generally fine with FL’s proposal. In addition, some comments as below:  **PSR/PDB**  Since there is a long gap between May meeting and Aug meeting, it’ll be good that RAN1#105-e can agree on some (PSR (Packet Success Rate), PDB) values, such guidance can help companies conduct related simulations during this long gap and good for progress.  Based on companies’ input, I/P-stream model have two characteristics:   * I-stream may have higher reliability than P-stream since I-stream is more important * P-stream may have smaller PDB since P-frame needs more time for decoding   To reflect these two characteristics above, and considering the baseline (PSR, PDB) for VR/AR single stream is (99%, 10ms), the following Proposal#1 is given:   * (PSR, PDB) values in row 1 and 2: this is to evaluate the case where I-stream has higher reliability than P-stream   + Row 1 refers to the case that I-stream has higher reliability than single-stream, and P-stream has lower reliability than single-stream.   + Row 2 refers to the case that I-stream has the same reliability as single-stream, and P-stream has lower reliability than single-stream.   + Note: the results for row 1 and 2 can be obtained via a single simulation since different PSR values only need some data processing work, so there is no additional simulation workload * (PSR, PDB) values in row 3: this is to evaluate the case where P-stream has smaller PDB than I-stream * (PSR, PDB) values in row 4, 5: these are to evaluate the case with larger and smaller PDB to reflect different network transmission quality, as RAN1#104b-e agreed for single stream as optional * The (PSR, PDB) values in row 1/2/3 are representative and can be taken as baselines. * The only difference between CG and AR/VR is the PDB of CG is relaxed by 5ms, as single stream does.   Btw: on reliability, single stream uses PSR (Packet Success Rate) X%. Instead of using PER, we suggest to use PSR or X or other names but with the same meaning for multi-stream too for easy discussion. Otherwise, sometimes we may use 99% (PSR), sometimes we may use 1% (PER), it may make discussions more complicated.  **Alpha**  On the value of alpha, we are open for both 1.5 and 2, and we also suggest to evaluate both values to gain more insight.  In summary, the following Proposal #1/#2 are suggested. If companies need more time to agree on these values, it might be ok to take Proposal#1,#2 as Working Assumptions so that RAN1 is still able to update these values in Aug meeting. Note that RAN1 has adopted similar method for single stream in previous meetings. In general, having the following Proposal#1/#2, rather than nothing, are good for progress since they provide some guidance for companies’ simulations before Aug meeting.  ***Proposal#1: For the optional two stream evaluation of I-frame and P-frame for DL video stream, the following multiple combinations of (PSR, PDB) are used for evaluation, other combinations of (PSR, PDB) can be optionally evaluated:***   |  |  |  | | --- | --- | --- | | ***Index*** | ***(Packet success rate X%, PDB (ms)) of {I-stream, P-stream}*** | | | ***VR/AR*** | ***CG*** | | *1 (baseline)* | *{(99.5, 10), (95, 10)}* | *{(99.5, 15), (95, 15)}* | | *2 (baseline)* | *{(99, 10), (90, 10)}* | *{(99, 15), (90, 15)}* | | *3 (baseline)* | *{(99, 17), (99, 9)}* | *{(99, 22), (99, 14)}* | | *4* | *{ (99.5, 7), (95, 7) }* | *{(99.5, 12), (95, 12)}* | | *5* | *{(95.5, 13), (90, 13)}* | *{(95.5, 18), (90, 18)}* |   ***Proposal#2:*** On the value of (average size ratio between one I-frame/slice and one P-frame/slice):   * Baseline: = [1.5 or 2] * Other values can be optionally evaluated, e.g. = 3   **Other comments**  We propose to also add “Urban macro” for FR1 since it’s also an important scenario for AR/VR in realistic deployment. Or we can simply remove the scenarios and leave it to company report, i.e. the same way as for single stream.  There is a typo in the row of “**Number of packets per stream at a time**” as below:   1. I-frame: 1 or 0   P-frame: 0 or 1 |
| Nokia, NSB | We also do not see the reason to model I and P frames separately, since it is not realistic to assume that both the gNB and the UE can always reliably distinguish between I and P frames. |

# DL: Two Stream Traffic Models: Option 2

Below is RAN1 agreement related to evaluation of two streams for DL.

|  |
| --- |
| RAN1 Agreement:  In addition to single stream per UE in DL which is baseline, two streams can be optionally evaluated for DL   * Option 1: I-frame + P-frame   + Option 1A: slice-based traffic model   + Option 1B: Group-Of-Picture (GOP) based traffic model * Option 2: video + audio/data * Option 3: FOV + omnidirectional stream * Companies should report detailed assumptions in their simulations on packet size distribution for each stream, packet arrival interval (or fps) for each stream, PDB for each stream, PER requirement for each stream, criteria for being satisfied. * Companies should strive to align the parameter values for the options chosen as much as possible * FFS: Whether audio stream is separate or aggregated with the data stream in option 2 (Intention of option 2 is not to create a 3 stream option) |

Company inputs on Option 2 are summarized below.

|  |  |
| --- | --- |
| **Company** | **Proposals in tdocs** |
| Apple | For DL traffic model Option 2, the audio/data flow is modeled with:   * A stream aggregating streams of audio and data   + Periodicity: 10ms   + Data rate: 0.756 Mbps/s or 1.12 Mbps   + Packet size: determined by periodicity and data rate   + PDB: 30 ms |
| LG | For optional two-stream DL traffic models, audio stream is aggregated with the data stream in Option 2. |
| DCM | *Audio stream is aggregated with data stream in option 2 for modeling of DL multiple streams* |

**Question 4. Please share your view on the above proposals for Option 2 from companies.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
| LG | We are not very supportive of modelling Option 2 itself. In this case, we don’t have a P-trace from SA4 with which we can try to align our model. Anyway, for the FFS part above, our understanding is that audio stream is aggregated with the data stream in Option 2. For traffic modelling of the audio/data stream in Option 2, we are okay with the Apple’s suggestion with the understanding that the model is based on SA4 input. |
| Apple | We propose to agree with modeling details for traffic models, the details are essentially the same as for uplink’s for AR2 data/audio stream. |
| QC | We are not very supportive of modelling Option 2 itself. Given the limited interest from companies, we do not see a strong motivation to agree on specific proposals. We want to leave detailed modeling as each companies’ assumption. |
| Samsung | We propose to leave the detailed modelling to companies. |
| vivo | We don’t see the need to model audio/data stream for XR evaluation. |
| CATT | We don’t think Option 2 would show any benefit or different in XR evaluation. |
| DOCOMO | We are fine with the Apple’s suggestion. |
| Apple-2 | Compared with Option 1 or Option 3, Option 2 is very simple. And Option 2 is clearly supported by the LS from SA4. From the agreement at #RAN1-104bis-e, agreeing on the details of Option 2 is important.   * “Companies should strive to align the parameter values for the options chosen as much as possible” |
| MTK | We do not see an obvious need to model audio/data stream for XR evaluation, but we are fine to accept Apple’s suggestion on including the detailed parameters (from SA4) to Option 2. |
| Ericsson | We do not see the need of a two-stream model. Agreeing on more optional simulation cases does not seem useful in light of the diverging simulation results. |
| ZTE, Sanechips | We are fine with Apple’s parameters for Option 2. In addition, according to S4aV200640, 1e-3 may be a reasonable PER for audio/data stream. |
| Nokia, NSB | The proposed values seem reasonable and was already agreed for UL AR optional stream for video+audio/data. |

# DL: Two Stream Traffic Models: Option 3

Below is RAN1 agreement related to evaluation of two streams for DL.

|  |
| --- |
| RAN1 Agreement:  In addition to single stream per UE in DL which is baseline, two streams can be optionally evaluated for DL   * Option 1: I-frame + P-frame   + Option 1A: slice-based traffic model   + Option 1B: Group-Of-Picture (GOP) based traffic model * Option 2: video + audio/data * Option 3: FOV + omnidirectional stream * Companies should report detailed assumptions in their simulations on packet size distribution for each stream, packet arrival interval (or fps) for each stream, PDB for each stream, PER requirement for each stream, criteria for being satisfied. * Companies should strive to align the parameter values for the options chosen as much as possible * FFS: Whether audio stream is separate or aggregated with the data stream in option 2 (Intention of option 2 is not to create a 3 stream option) |

Company inputs on Option 3 are summarized below.

|  |  |
| --- | --- |
| **Company** | **Proposals in tdocs** |
| MTK | ***Proposal 6: Discuss two possible options:***   1. ***FoV vs. non-FoV*** 2. ***FoV vs. low resolution Omnidirectional stream***  * ***FFS: Need for different QoS requirements for the two streams.***   ***FFS: co-existence with the QoS requirements for I/P-frames.*** |
| ZTE | ***Further discussion in RAN1 the parameters of FoV and non-FoV stream modelling for DL 360°video stream and parameters in table 8 can be regarded as starting point***  ***Table 8 Initial Parameters of FoV and non-FoV stream modelling***   |  |  |  | | --- | --- | --- | | ***Application*** | ***VR1*** | | | ***Two Stream Data*** | ***Stream #1: FoV stream***  ***Stream #2: Non-FoV stream*** | | | ***Option 1: sliced based traffic model*** | ***Option 2: Two separate streams*** | | ***Structure*** | ***A frame consists of:***  ***Stream #1: 1 (18 tiles)***  ***Stream #2: 1*** | ***A Group of Tiles consist of: Stream #1: 18 tiles***  ***Stream #2: 1*** | | ***Frame Per Second*** | ***Stream #1: 30FPS***  ***Stream #2: 30FPS*** | ***Stream #1: 540 tiles per second***  ***Stream #2: 30FPS*** | |

**Question 5. Please share your view on the above proposals for Option 3 from companies.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
| LG | Similar comment as above. For Option 3, we also don’t have a P-trace from SA4 with which we can try to align our model. We prefer not to spend time on discussing details of parameter values until we have a clear guidance from SA4. |
| QC | We are not very supportive of modelling Option 3 itself. Given the limited interest from companies, we do not see a strong motivation to agree on specific proposals. We want to leave detailed modeling as each companies’ assumption. |
| Samsung | Same as above. We propose to leave the detailed modelling to companies. |
| vivo | It is still unclear how to model multi-stream by FOV + omnidirectional streams. More input on the detailed modelling is needed. We don’t have strong preference to adopt option 3. If companies have interest on evaluating option 3 optionally, the details can be up to company report. |
| CATT | We are not supportive of Option 3. |
| MTK | The concept of group of tile is interesting but new to us. We are open to discuss more details on Option 3 if companies have interest. |
| Ericsson | We do not see the need of a two-stream model. Agreeing on more optional simulation cases does not seem useful in light of the diverging simulation results. |
| ZTE, Sanechips | We are fine with Option 3. For Option 3, FoV and non-FoV traffic model is distinctive of the VR1 traffic. And some reference parameters including, e.g. data rate for FoV/non-FoV stream, frame per second and E2E downlink budget, are provided in S4aV200614. If companies are interested in Option 3, we are able to ask SA4 for the KPIs of FoV/non-FoV streams, or reuse the KPIs of I/P-streams as that of FoV/non-FoV streams. |
| Huawei, HiSilicon | We are open to discuss Option 3.  In the recent SA4 LS to RAN1 (R1-2104023), SA4 mentioned they are improving traffic models for VR1 (cyan part below). So we expect there will be more progress in SA4 on VR1 in the near future.   |  | | --- | | … 1 Overall description 3GPP TSG SA WG4 would like to inform 3GPP TSG RAN WG1 on our progress on XR Traffic for the purpose of evaluating the performance of XR application and traffic on 5G Systems and in particular the new radio.  For the study item, we envision completion date in September 2021. All information is collected in the Permanent Document (as attached). We will move all agreements for proper documentation in a newly assigned TR 26.926.  We have progressed work on the following topics since our latest communication:   * Improved traffic models for VR1 * Usage of Traces to develop statistical models   … | |
| Nokia, NSB | We also do not support modelling this option. We do not see how we converge with the results if companies, interested in modelling, will have their own assumptions related to that Option. Even with agreed baseline model we are not well aligned, and it is hard to make final conclusions. |

# Common baseline

Given that there are a number of simulation scenarios, i.e., combinations of applications, data rates, PDB, etc., it is challenging to compare simulation results among companies. In this regard, the following scenarios are proposed in [18] as a common baseline.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Data rate  [Mbps] | Frame rate  [FPS] | PDB  [ms] | Jitter? | Random frame size? |
| DL | CG | 8 | 60 | 15 | Yes | Yes |
|  | AR/VR | 30 | 60 | 10 | Yes | Yes |
| UL | Pose | 0.2 | 250 | 10 | No | No |
|  | Scene | 10 | 60 | 60 | No | Yes |

Table 1 The traffic parameters of the common baseline

* RAN1 should use the traffic model settings in Table 1 as a common baseline.
* In the common baseline, 99% of the frames should arrive within the PDB
* For the capacity simulations in the common baseline, only one stream at a time is simulated.

**Question 6. Please share your view on the above proposal, i.e., defining those scenarios/parameters as the common baseline. It is clarified that the potential definition of such a common baseline set is not intended for prioritization of simulation scenarios but is purely intended to facilitate comparison of results among companies by ensuring that we have more results from more companies for the reduced set of scenarios. With a potential agreement on the common baseline, companies are generally encouraged to submit results for other baseline and optional scenarios/configurations as many as they can.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
| LG | We are supportive of the intention of setting a common baseline in principle. We are open to discuss details. |
| Apple | The intention from Ericsson [18] is very good, but it may actually add to the load of evaluation. Note for calibration, we are open to very simple setup to check alignment among companies. But the discussion on calibration and capacity evaluation should be separated. |
| QC | We support the FL proposal. |
| Samsung | We support the idea of a common agreed set of well-identified XR simulation scenarios. But in absence of some calibration phase, we think that even for this narrowed-down set of simulation cases, we will still observe a lot of variance in company reported results. |
| InterDigital | We are ok in principle to apply a common baseline for showing the results for UL and DL for CG/VR/AR as it would facilitate clearer comparison of results among companies as well as comparison of gain/loss of other optional results of a company with respect to its baseline. Whether the parameters and notes in Table 1 above can be used as common baseline should be further discussed. |
| Vivo | We support the intention of setting a common baseline for comparison of results among companies.  Besides, it may need to clarify that it is not necessary to evaluate these common baseline scenarios for all the deployments, i.e. DU/InH/Uma. Company can choose the deployment for the common baseline. |
| CATT | We are OK to have the baseline configuration for calibration and comparison. |
| DOCOMO | We support the FL proposal. |
| MTK | We support the FL proposal. |
| Ericsson | We support the FL proposal. |
| Xiaomi | We support the FL proposal. |
| ZTE, Sanechips | We agree with FL proposal. |
| Huawei, HiSilicon | Since RAN1 already agreed baselines for different parameters in separate agreements, we are not very clear about the necessity of this additional agreement. If the common baseline is the same as previous agreements, then there is no need to have it. If not, then are we going to revert previous agreements?  Btw: AR UL PDB is still under discussion, so the discussion of this issue can be delayed a little bit. |
| Nokia, NSB | We generally support to have common parameters to simulate. However, we suggest changing the data rate for CG to higher value – 30 Mbit/s. The reason is that 8 Mbit/s does not seem to be a challenging case itself. Many companies did not reach a final number of users that can be supported since it is a lot of users to model. This in turn increases the simulation time when trying to reach the final number of UEs. Therefore, it might not be the best option to put it as a baseline. |

# Summary of Proposals in Tdocs [1-18]

|  |  |
| --- | --- |
| **Company** | **Proposals in tdocs** |
| FUTUREWEI [1] | ***For remaining parameters in [] for Stream 2 in Option 1 and 3 and [] for the stream in Option 2, suggest that the 60msec is taken as baseline and remove the []. For optional PDB, suggest down-selection between 10 msec or 15 msec*** |
| Huawei [2] | ***Proposal 1: For video of AR/VR/CG, adopt the traffic model in following Table 5 for “Option 1: I-frame + P-frame”.***  Table 5. Option 1: I-frame + P-frame model for DL video   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Two data streams, i.e. M1 = 2** | **Option 1A: slice-based** | | | **Option 1B: GOP-based** | | | | I-stream | | P-stream | I-stream | | P-stream | | **Packet modelling** | Slice-level | | | Frame-level | | | | **Traffic pattern** | Both streams are periodic with the same FPS. | | | Follow the GOP structure, e.g. GOP size K = 8. | | | | **Number of packets per stream at a time** | 1 | | N-1 | 1 or 0 | | 0 or 1 | | * N is the number of slice per frame, e.g. N = 8. | | | | **Average data rate per stream** |  | |  |  | |  | | * R: average data rate of a single stream video * : average size ratio between one I-frame/slice and one P-frame/slice, e.g. = 1.5, 2   + Other values can be optionally evaluated. | | | | | | | **Packet size distribution** | truncated Gaussian distribution | | | | | | | Mean = | Mean = | | Mean = | Mean = | | | * [STD, Max, Min]: [10.5, 150, 50]% of Mean packet size * FPS is the frame rate of the single stream video | | | | | |   Note: the QoS requirement for each stream is separately discussed in the KPI part. |
| vivo [3] | *Proposal 1: For dual-eye buffer, the dual-eye buffer traffic model in Table 1 can be optionally evaluated.*  *Proposal 2: Which traffic model to be chosen as the baseline of multi-stream evaluation needs to be further studied.*  *Proposal 3: For multi-stream modelling with I-frame and P-frame, the following traffic characteristics need to be further studied,*   * *The packet size ratio of I-frame and P-frame.* * *The length of GOP for GOP-based model and the number of slices per frame for Slice-based model.* * *The jitter distribution.*   *Proposal 4: How to set the X value and air interface PDB requirements for multi-stream model need to be further studied.*  ***Proposal 5: A UE with multi-stream is declared as a satisfied UE if each stream from the multi-stream has been satisfied, i.e. for each stream more than X (%) of packets are successfully transmitted within a given air interface PDB, where the X value and the given air interface PDB can be set per stream.*** |
| CATT [4] | ***Proposal 1: When single eye or dual eye buffer is modelled for evaluation, the total transmission data rate would be kept the same with the details of modelling reported by companies.***  ***Proposal 2: I-frame and P-frame would not be modelled differently as the baseline.***  ***Proposal 3: The audio stream could be aggregated with the data stream in option 2 (video + audio/data) for two streams per UE in DL.***  ***Proposal 4: The PDB for a stream aggregating streams of scene, video, data, and audio would be either 10ms or 15ms.*** |
| Nokia [5] | **Observation 1:** *Table 1, 2, and 3 are the summary of the main parameters agreed for the DL and UL baseline traffic models and employed for the evaluation purposes.*  **Observation 2:** *Table 4 is the summary of the mandatory streams to be used for evaluation of VR, AR, and CG. Table 5 shows the optional streams that can be also modelled when evaluating VR, AR, and CG applications.* |
| Qualcomm [6] | ***Proposal: For UL AR traffic model, conform PDB values in [ ] for Stream 2 in Option 1 and 3, and Option 2, i.e., 60 ms PDB for the stream representing streams of video, camera feed, audio, and data*** |
| OPPO [7] | ***For evaluations of AR in UL:***   * ***Option 1 (Baseline for power and capacity evaluations): Two streams as defined below***    + ***Stream 1: pose/control***     - ***Traffic model and QoS parameters are same as for pose/control for UL CG/VR.***   + ***Stream 2: A stream aggregating streams of scene, video, data, and audio.***      - ***Packet size: Truncated Gaussian distribution with the parameter values same as for DL***     - ***Periodicity: 60 fps***       * ***Jitter (optional): same model as for DL***     - ***Data rate: 10 Mbps (baseline), 20 Mbps (optional)***     - ***PDB: ~~[~~15~~60]~~ ms (baseline), ~~[~~10~~/15]~~ ms (optional)*** * ***Option 2 (Optional for power evaluation and baseline for capacity evaluation): Single stream as defined below***    + ***Packet size: Truncated Gaussian distribution with the parameter values same as for DL***   + ***Periodicity: 60 fps***     - ***Jitter (optional): same model as for DL***   + ***Data rate: 10 Mbps (baseline), 20 Mbps (optional)***   + ***PDB: ~~[~~15~~60]~~ ms (baseline), ~~[~~10~~/15]~~ ms (optional)*** * ***Option 3 (Optional): Three streams as defined below***    + ***Stream 1: pose/control***     - ***Traffic model and QoS parameters are same as for pose/control for UL CG/VR.***   + ***Stream 2: A stream aggregating streams of scene and video***      - ***Packet size: Truncated Gaussian distribution with the parameter values same as for DL***     - ***Periodicity: 60 fps***       * ***Jitter (optional): same model as for DL***     - ***Data rate: 10 Mbps (baseline), 20 Mbps (optional)***     - ***PDB: ~~[~~15~~60]~~ ms (baseline), ~~[~~10~~/15]~~ ms (optional)***   + ***Stream 3: A stream aggregating streams of audio and data***      - ***Periodicity: 10ms***     - ***Data rate: 0.756 Mbps/s or 1.12 Mbps***     - ***Packet size: determined by periodicity and data rate***     - ***PDB: 30 ms***   Proposal 2: Modify the following RAN1 agreement of RAN#104bis-e as below (changes are highlighted by RED)  ***On UL Traffic model and QoS parameters***   * ***CG/VR: single stream (pose/control)*** * ***Traffic model for Pose/control***    + ***Periodic: 4ms (no jitter)***      - ***Other values can be optionally evaluated.***   + ***Fixed: 100 bytes***      - ***PDB: 4ms (baseline), 10 ms (Optional).***   + ***A UE is declared a satisfied UE if more than X (%) of packets are successfully delivered within the given air interface PDB.***      - ***The baseline X value is 99.***     - ***Other X values can be optionally evaluated are 90 and 95.***   Proposal 3: For air interface PDB for DL video stream, no more mandatory value is needed. |
| Intel [8] | ***Based on SA4 discussions for XR conversational, 60 ms uplink PDB seems reasonable (for a E2E latency budget of 100ms - 200ms)*** |
| Apple [9] | **Proposal 1:**  For DL traffic model Option 2, the audio/data flow is modeled with:   * A stream aggregating streams of audio and data   + Periodicity: 10ms   + Data rate: 0.756 Mbps/s or 1.12 Mbps   + Packet size: determined by periodicity and data rate   + PDB: 30 ms   **Proposal 2**: The Packet delay budget for the video stream in UL traffic models Option 1/2/3 is 10 ms. |
| Sony [10] | **Proposal 1: Reuse the media characteristics / simulation configurations that SA4 already defined in LS - R1-2101765 [4].**  **Proposal 2: RAN1 strives to define one or more scenarios that can be used for simulation calibration purposes.**  **Proposal 3: Consider the entire video stream (I-frames, P-frames etc) to be transported on a bearer with a single associated QoS class.**  **Proposal 4: RAN1 to study layer-1 aspects of large packet transmission with better reliability than eMBB and/or with low packet delay** |
| Samsung [11] | **Proposal 1**  *The dual-eye buffer case is evaluated as single stream case by doubling the FPS and halving the packet size compared to the single-eye buffer case.*  **Proposal 2**  *PDB 60 ms (baseline), 10/15 ms (optional) for Option 1 & 3 Stream 2 and for Option 2* |
| MediaTek  [12] | ***Proposal 1: For the optionally evaluated two-stream DL traffic, prioritize the IDR Group-Of-Picture (GOP) based traffic model, Option 1B) evaluation.***  ***Proposal 2: For the optionally evaluated two-stream DL traffic agreed in RAN1 #104-bis-e, further agree the baseline evaluation parameters about traffic model for Option 1 as in Table I.***  **Table I**   |  |  |  | | --- | --- | --- | | Application | AR/VR/CG | | | Two data streams, i.e. M1 = 2 | * Stream #1: I-stream * Stream #2: P-stream | | | Option 1: slice-based | Option 2: frame-based | | Packet modelling | Slice-level | Frame-level | | Packet size distribution | Truncated Gaussian distribution for each stream. The STD/Max/Min of packet size follows the single stream assumption: [STD, Max, Min]: [10.5, 150, 50]% of Mean packet size | | | Mean packet size | Derived from FPS and average data rate listed below | | | Traffic arrival pattern | Both streams are periodic with 60 FPS. | Follow the GOP structure with GOP size K = 8, i.e., one I-frame is followed by seven P-frames. The overall FPS of I-frame plus P-frame is 60. | | Number of packets per stream at a time | * Stream #1: 1 * Stream #2: N-1 * N is the number of slice per frame, use N = 8. | * Stream #1: 1 or 0 * Stream #2: 0 or 1 * At each time instant, there is either only one I-stream packet or only one P-stream packet | | Average data rate | Stream #1: Stream #2 = : (N-1)  The aggregated data rate of I-stream plus P-stream should be the same as single stream assumptions | Stream #1: Stream #2 = : (K-1)  The aggregated data rate of I-stream plus P-stream should be the same as single stream assumptions | | is average size ratio between one I-frame/slice and one P-frame/slice, use = 3   * Other values can be optionally evaluated. | |   ***Proposal 3: Evaluate [PER\_I, PER\_P] = [1%, 1%] and [0.1%, 5%].***  ***Proposal 4: Evaluate [PDB\_I, PDB\_P] = [10ms, 10ms] and [17ms, 9ms]. It is noted that for GOP size K=8 in Option 1B, [PDB\_I, PDB\_P] = [10ms, 10ms] and [17ms, 9ms] provide the same average PDB.***  ***Proposal 5: RAN1 to coordinate and cooperate with SA4 to construct a video quality evaluation block (as shown in the red block in Figure 3 below) based on statistical models used in RAN1 to evaluate the different QoS requirements and the performance enhancement for various RAN1 proposals.***  ***Proposal 6: Discuss two possible options:***   1. ***FoV vs. non-FoV*** 2. ***FoV vs. low resolution Omnidirectional stream***  * ***FFS: Need for different QoS requirements for the two streams.*** * ***FFS: co-existence with the QoS requirements for I/P-frames.*** |
| LG [13] | *Proposal 1: Single and dual eye buffer can be modelled using the baseline single-stream DL traffic model with the agreed baseline and optional parameters.*   * *No further discussion on details is needed.*   ***Proposal 2: Companies to report the X value used for per UE KPI including multiple X values if they used.***   * *No further discussion on the X values is needed.*   ***Proposal 3: For evaluations of AR in UL, assume the following PDB values for Stream 2 in Option 1 and 3, and Option 2***   * ***PDB: 10/15 ms (baseline), 60 ms (optional)***   ***Proposal 4: For optional two-stream DL traffic models, audio stream is aggregated with the data stream in Option 2*** |
| InterDigital [14] | **Observation 1:** For CG, the UL traffic can be represented by 2 different traffic streams delivering the following data:   * User actions (e.g. gamepad controller, HMD): traffic volume (pkts/s) follows user actions and stops when there is no activity * Control data (e.g. protocol flow control, keep-alive messages):   + Reports: traffic volume is correlated to DL traffic and is independent of the user activity   + Lifecycle status: traffic volume is constant and periodic (e.g. keep-alive every 500ms)   **Observation 2:** For CG, the chacteristics of the UL traffic streams are as follows:   * User Actions   + Packet arrival is aperiodic and correlated with user activity. Inter-packet arrival follows exponential distribution   + Packet size can vary depending on the CG platform (i.e. typically Gaussian distribution) * Control Data   + Packet arrival can be periodic (e.g. 1/100Hz, 1/10Hz) or correlated with user activity depending on the CG platform. Dedicated periodic flows may exist for indicating lifecycle status.   + Packet size can vary depending on the CG platform (i.e. typically Gaussian distribution)   **Observation 3:** For CG, the different UL traffic streams (e.g. user actions, control data) are delivered using different transport protocols  **Observation 4:** For a given application, using multiple traffic streams with different characteristics can result in impacting both capacity and power consumption performance  **Observation 5:** The sensitivity of QoE to change in QoS is significantly different between the 2 traffic streams belonging to the same CG application  **Observation 6:** The traffic models applied for AR/VR use case (TR 38.824) and UL pose/control traffic can be used to model the 2 UL traffic streams related to CG (i.e. user action and control data streams)  **Observation 7:** Similar to CG, VR uses 2 traffic streams in UL with similar traffic characteristics  **Observation 8:** For AR, the UL traffic characteristics of aggregated video/media stream is significantly different than the pose/control stream  **Observation 9:** For AR, due to the per-UE KPI requiring equal importance for all streams for meeting the X% and PDB and pairing of 2 streams with significant differences, the pose/control stream can have major impact on the capacity achievable  Based on these observations, the following conclusions were made:  **Proposal 1:** RAN1 uses 2 traffic streams in UL (i.e. user actions and control data) for CG as baseline for capacity and power consumption evaluations  **Proposal 2:** RAN1 uses the following traffic models in UL for evaluating 2 streams for CG   * Stream 1: User actions   + - Periodic with different (inter-packet) arrival time: 10ms (average)     - Fixed: 200 bytes     - PDB: 4 ms * Stream 2: Control data   + - Periodic: 4ms (no jitter)     - Fixed: 100 bytes     - PDB: 10 ms   **Proposal 3:** RAN1 uses 2 traffic streams in UL (i.e. user actions and pose/control data) for VR as baseline for capacity and power consumption evaluations  **Proposal 4:** VR evaluations in UL use the same traffic models used for CG for the 2 traffic streams (i.e. user actions and pose/control data)  **Proposal 5:** RAN1 uses the 2 traffic streams in UL (i.e. pose/control and aggregated video/media) for AR as baseline for capacity evaluations  **Proposal 6:** RAN1 uses the following traffic model in UL as baseline for evaluating the aggregated traffic stream (i.e. consisting of scene, video, data, and audio) for non-conversationalAR:   * + - Packet size: Truncated Gaussian distribution with the parameter values same as for DL     - Periodicity: 60 fps       * Jitter (optional): same model as for DL     - Data rate: 10 Mbps (baseline), 20 Mbps (optional)     - PDB: [10] ms (baseline) |
| Xiaomi [15] | **Proposal 1: For XR DL evaluation, interleaved eye buffer model can be optionally considered.**  **Proposal 2: Send LS to SA4 to ask (X, PDB) requirement and statistical models for packets associated to I-frames and P-frames** |
| ZTE [16] | [*Observation 1:* There is maximum 7% capacity gain for single eye traffic in high system load , compared to dual eye traffic in indoor house scenario, when bit rate is 45Mbps.](#_Toc15034)  [*Proposal 1:* The simulation results based on traffic for dual eye buffer is enough to present the system capacity. And companies can provide the results based on traffic for single eye buffer optionally.](#_Toc9207)  [*Proposal 2:* Confirm the note on 60ms as baseline for the PDB of stream including video. Other reasonable PDB values can be evaluated optionally.](#_Toc22903)  [*Proposal 3:* Further discussion in RAN1 the parameters of I/P stream modelling for DL video stream and parameters in table 7 can be regarded as starting point.](#_Toc26894)  ***Table 7 Summary of parameters for I/P stream modelling***   |  |  |  | | --- | --- | --- | | ***Application*** | ***AR/VR/CG*** | | | ***Two stream data*** | ***Stream #1: I-frame***  ***Stream #2: P-frame*** | | | ***Option 1 Sliced-based*** | ***Option 2: Frame-based (GoP)*** | | ***Structure*** | ***A frame consists of:***  ***Number of Stream #1: 1***  ***Number of Stream #2: 7*** | ***A GoP consists of:***  ***Number of Stream #1: 1***  ***Number of Stream #2: 7*** | | ***Frame per second*** | ***Stream #1: 60FPS***  ***Stream #2: 60FPS*** | ***Stream #1 + Stream #2 = 60FPS*** | | ***Average packet size ratio*** | ***Stream #1 : Stream #2 = 2:1*** | ***Stream #1 : Stream #2 = 2:1*** | | ***(PSR, PDB)*** | ***AR/VR:***  ***Stream #1: (99%, 20ms)***  ***Stream #2: (90%, 20ms)***  ***CG: Stream #1: (99%, 20ms)***  ***Stream #2: (90%, 20ms)*** | ***Option 1:***  ***Stream #1: (99%, 10ms)***  ***Stream #2: (90%, 10ms)***  ***Option 2:***  ***Stream #1: (99%, 15ms)***  ***Stream #2: (99%, 10ms)*** |   [*Proposal 4:* Further discussion in RAN1 the parameters of FoV and non-FoV stream modelling for DL 360°video stream and parameters in table 8 can be regarded as starting point.](#_Toc31285)  ***Table 8 Initial Parameters of FoV and non-FoV stream modelling***   |  |  |  | | --- | --- | --- | | ***Application*** | ***VR1*** | | | ***Two Stream Data*** | ***Stream #1: FoV stream***  ***Stream #2: Non-FoV stream*** | | | ***Option 1: sliced based traffic model*** | ***Option 2: Two separate streams*** | | ***Structure*** | ***A frame consists of:***  ***Stream #1: 1 (18 tiles)***  ***Stream #2: 1*** | ***A Group of Tiles consist of: Stream #1: 18 tiles***  ***Stream #2: 1*** | | ***Frame Per Second*** | ***Stream #1: 30FPS***  ***Stream #2: 30FPS*** | ***Stream #1: 540 tiles per second***  ***Stream #2: 30FPS*** | | ***Data Rate*** | ***Stream #1: 12.78 Mbps***  ***Stream #2: 8Mbps*** | ***Stream #1: 12.78Mbps (the aggregated data rate of the 18 tiles within a group of tiles)***  ***Stream #2: 8Mbps*** | | ***(PSR, PDB)*** | ***Stream #1: (99%, 20ms)***  ***Stream #2: (90%, 20ms)*** | ***Stream #1: (99%, 10ms)***  ***Stream #2: (90%, 10ms)*** | |
| DOCOMO [17] | **Proposal 1:**   * *Consider to study XR conference as optional.*   **Proposal 2:**   * *Audio stream is aggregated with data stream in option 2 for modeling of DL multiple streams.*   **Proposal 3:**   * *10 or 15 ms should be baseline for PDB of AR in UL.* |
| Ericsson [18] | [Observation 1 Multiple options even for baseline cases have been agreed and complexity of evaluations has increased](#_Toc71642547)  [Observation 2 Many options and high complexity of simulation settings might be problematic for comparing and understanding results between companies](#_Toc71642548)  [Observation 3 Understanding base settings is of high importance to reach the goal of the XR SI](#_Toc71642549)  [Observation 4 Parameters to decide on consists of, UL/DL traffic types, data rates, frame rates, PDB, percentile of packets to meet PDB, number of simultaneous streams and jitter](#_Toc71642550)  [Proposal 1 Define a common baseline for all companies to use in their evaluations](#_Toc71642551)  [Proposal 2 RAN1 should use the traffic model settings in Table 1 as a common baseline.](#_Toc71642552)  [Proposal 3 In the common baseline, 99% of the frames should arrive within the PDB](#_Toc71642553)  [Proposal 4 For the capacity simulations in the common baseline, only one stream at a time is simulated](#_Toc71642554) |

# List of contributions in RAN1 #105-e

1. [R1-2104207](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104207.zip) XR traffic model FUTUREWEI
2. [R1-2104238](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104238.zip) Traffic model for XR and Cloud Gaming Huawei, HiSilicon
3. [R1-2104395](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104395.zip) Remaining issues on traffic models of XR vivo
4. [R1-2104502](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104502.zip) XR traffic model CATT
5. [R1-2104555](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104555.zip) On Traffic Model for XR study Nokia, Nokia Shanghai Bell
6. [R1-2104701](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104701.zip) Remaining Issues on XR Traffic Models Qualcomm Incorporated
7. [R1-2104745](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104745.zip) Discussion on the XR traffic models for evaluation OPPO
8. [R1-2104934](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2104934.zip) Traffic Model for XR Intel Corporation
9. [R1-2105134](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105134.zip) Considerartions on XR traffic model Apple
10. [R1-2105181](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105181.zip) Considerations on XR traffic model Sony
11. [R1-2105342](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105342.zip) Traffic Models for XR Samsung
12. [R1-2105376](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105376.zip) Traffic Model for XR and CG MediaTek Inc.
13. [R1-2105443](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105443.zip) Discussion on traffic models for XR evaluation LG Electronics
14. [R1-2105499](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105499.zip) Discussion on UL traffic models InterDigital, Inc.
15. [R1-2105547](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105547.zip) Discussion on remaining issues of traffic Model for XR services Xiaomi
16. [R1-2105603](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105603.zip) Remaining Issues of XR Traffic Model ZTE, Sanechips
17. [R1-2105726](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105726.zip) Discussion on traffic model for XR NTT DOCOMO, INC.
18. [R1-2105829](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_105\Docs\R1-2105829.zip) Traffic model for XR Ericsson