3GPP TSG-RAN WG1 Meeting #104-e R1-21xxxxx

e-Meeting, April 12th – 20th, 2021

Source: Moderator (Qualcomm)

Title: Email discussion for XR traffic models and KPIs

Agenda Item: 8.14.1

Document for: Discussion and Decision

# Introduction

This contribution is a summary on the email discussion on remaining open issues for traffic models and KPI’s for XR and Cloud Gaming.

[104b-e-NR-XR-01] Email discussion/approval on traffic model – Eddy (Qualcomm)

* 1st check point: April 15
* Final check point: April 20

# Outcomes of RAN1 #104b-e

# Discussion

## DL Traffic Model: Single Stream

1. **DL packet size distribution**

RAN#104-e Working assumption

|  |
| --- |
| (Working assumption) Parameters of Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)   * 1. Mean: Derived from average data rate and fps as follows.      1. (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]   2. STD      1. TBD   3. Max packet size      1. TBD   4. Min packet size      1. TBD      2. FFS whether or not to use this parameter |

Companies’ views in RAN1#104bis-e tdocs are presented in the table below.

|  |  |
| --- | --- |
| Huawei | STD: 15% of Mean packet size  Max packet size: 2 \* Mean packet size  Min packet size : 25% of Mean packet size |
| OPPO | For the distribution of packet sizes, confirm the working assumption of truncated Gaussian distribution by removing the brackets, i.e.,   * STD: 15% or 20% of Mean packet size * Max packet size: 1.5 x Mean packet size * Min packet size: 0.5 x Mean packet size |
| vivo | * STD   + 15% of Mean packet size derived above * Max packet size   + 1.5 x Mean packet size derived above * Min packet size   + 0.1 \* Mean packet size derived above |
| CATT | The truncated Gaussian distribution for VBR  Uniform distribution for CBR.  For the parameters of the statistical distribution for Packet size   * STD: 5% and 10% of Mean packet size for CBR and VBR, respectively. * Max: 1.2 time of and equal to Mean packet size for VBR and CBR, respectively. * Min: limited by the minimum IP packet size, i.e. 46Bytes. |
| MTK | * STD: 15% of Mean packet size derived above * Range: 1.5 × Mean packet size derived above |
| Futurewei | * STD: 15% of Mean packet size * Max packet size: 1.5 x Mean packet size * Min packet size: 0.5 x Mean packet size |
| Nokia | * STD: 2% of mean packet (frame) size * Max packet (frame) size: 1.1 x mean packet (frame) size * Min packet (frame) size: 0.9 x mean packet (frame) size |
| Ericsson | * STD: 15% of mean packet (frame) size * Max packet (frame) size: 1.5 x mean packet (frame) size * Min packet (frame) size: 0.5 x mean packet (frame) size |
| Intel | * for CBR configurations the frame-size variations are quite small   + the max/mean frame-size ratio is ~ 1.06   + the min/mean frame-size ratio is ~0.93   + the std/mean frame-size ratio is ~0.02 * for cVBR configurations the frame-size variations are larger   + the max/mean frame-size ratio is ~ 1.18 – 1.94 with smaller ratios corresponding to 8 slice/eye buffer case while large ratio corresponding to 1 slice case   + the min/mean frame-size ratio is ~ 0.24 – 0.48 with larger ratios corresponding to 8 slice/eye buffer case while small ratio corresponding to 1 slice case   + the std/mean frame-size ratio is ~ 0.07 – 0.14 |
| Qualcomm | * STD: 7.5% of Mean * Max packet size: 135% of Mean * Min packet size: 54.5% of Mean |
| Samsung | * STD: 15% of Mean * Max packet size: 150% of Mean |
| ZTE | * [Single eye packet size](#_Toc68618183)   + [STD = 4% \* mean, MAX = 112% \* mean](#_Toc68618184) * [Dual eye packet size](#_Toc68618185)   + [STD = 3% \* mean, MAX = 109% \* mean.](#_Toc68618186) |

**Summary**:

* From SA4 trace files: STD, Max, and Min depend on CBR, VBR, single vs. multiple (8) slices per frame, etc.
* Two companies address different values for CBR and VBR. However, moderator recommends to not evaluate separately CBR and VBR based on majority view.
* Average values from companies’ inputs
  + STD: 10.5% of Mean packet size
  + Max: 135% of Mean packet size
  + Min: 50% of Mean packet size

1. **Based on the discussions and proposals in tdocs for RAN1#104bis-e, the moderator makes the following proposal. Please share your comments.**

**Moderator proposal**: 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)

* Mean packet size: Derived from average data rate and fps as follows.
  + (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
* STD: 10.5% of Mean packet size
* Max: 135% of Mean packet size
* Min: 50% of Mean packet size

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

1. **DL Jitter Model**

RAN1#104-e agreement on Jitter for DL video stream for a single UE

|  |
| --- |
| * Per the agreed statistical traffic model, arrival time of packet k is k/X1000 [ms] + J [ms], where X is the given fps value and J is a random variable.  * J is drawn from a truncated Gaussian distribution:   + Mean: [0]   + STD: [2 ms]   + Range: [[-4, 4]ms]     - Note: The values ensure that packet arrivals are in order (i.e., arrival time of a next packet is always larger than that of the previous packet)   + Note: The above values for mean, STD and Range are working assumption for initial simulations, and is to be revisited potentially with more inputs from companies in RAN1#104-bis-e |

Companies’ views in RAN1#104bis-e tdocs are presented in the table below.

|  |  |
| --- | --- |
| Huawei | Mean: 0; STD: 2 ms; Range: [-4, 4]ms |
| OPPO | Mean: 0; STD: 3 ms; Range: [-5, 5]ms |
| vivo | Mean: 0; STD: 2 ms; Range: [-4, 4]ms |
| CATT | Mean: 20 ms; STD: 6.35 ms; Range: [9, 31] ms |
| MTK | Mean: 0; STD: 2 ms; Range: [-4, 4]ms |
| Nokia | Mean: 0; STD: 3 ms; Range: [-6, 6]ms |
| Ericsson | Mean: 0; STD: 2 ms; Range: [-4, 4]ms |
| Qualcomm | Mean: 0; STD: 2 ms; Range: [-4, 4]ms |
| ZTE | Presented different values for different VR2 configurations. |

**Summary**:

* The clear majority of companies propose to confirm the RAN1#104-e Working Assumption, Mean: 0; STD: 2 ms; Range: [-4, 4]ms.

1. **Based on the discussions and proposals in tdocs for RAN1#104bis-e, the moderator makes the following proposal. Please share your comments.**

**Moderator proposal:** Confirm the following WA.

* Jitter for DL video stream for a single UE
  + Per the agreed statistical traffic model, arrival time of packet k is k/X x 1000 [ms] + J [ms], where X is the given fps value and J is a random variable.
  + J is drawn from a truncated Gaussian distribution:
    - Mean: 0 ms
    - STD: 2 ms
    - Range: [4, 4] ms
      * Note: The values ensure that packet arrivals are in order (i.e., arrival time of a next packet is always larger than that of the previous packet)
    - Other values can be optionally evaluated

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

1. **DL Per UE KPI (Baseline): Definition of whether each UE is satisfied or not in case of single DL stream per UE.**

RAN1#104-e agreement

|  |
| --- |
| * 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, e.g., 99, 95     - FFS different values for I-frame and P-frame if evaluation of them is agreed.   + Other values can be optionally evaluated |

Companies’ views in RAN1#104bis-e tdocs are presented in the table below.

|  |  |
| --- | --- |
| Huawei | X=99, 95, <95 |
| OPPO | 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} |
| vivo | 99 |
| MTK | 99 |
| Nokia | 99 |
| Ericsson | 99 |
| Xiaomi | 99.9 |
| Qualcomm | 99 |
| Samsung | 99 |
| ZTE | 99 |
| LGE | 95 |
| InterDigital | For a given XR/CG applications, employ link performance metrics such as user DL/UL throughput to determine a user’s experience satisfaction by the percentage (70%, 90%) of achieved maximum throughput |

**Summary**:

* X = 99 (7 companies)
* X = 95 (1 company)
* X = 99.9 (1 company)
* There are some other views.

1. **Based on the discussions and proposals in tdocs for RAN1#104bis-e, the moderator makes the following proposal. Please share your comments.**

**Moderator proposal:**

* In case of single stream per UE in DL, a UE is declared a satisfied UE if more than X (%) of packets are successfully delivered within a given air interface PDB.
  + The baseline X value is 99.
    - Other values can be optionally evaluated
    - FFS different values for I-frame and P-frame if evaluation of them is agreed.

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

1. **DL Per UE KPI (Optional): Definition of whether each UE is satisfied or not in case of single DL stream per UE.**

RAN1#104-e agreement

|  |
| --- |
| * Baseline: A UE is declared a satisfied UE if more than X (%) of packets are successfully delivered within a given air interface PDB.   + The exact value of X is FFS, e.g., 99, 95     - FFS different values for I-frame and P-frame if evaluation of them is agreed.   + Other values can be optionally evaluated * Per UE KPI   + 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 |

Companies’ views in RAN1#104bis-e tdocs are presented in the table below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Huawei | Proposal 2: XR Quality Index (XQI) is defined to reflect the impact of network transmission on user experience in XR and CG services.  Observation 1: In real XR/CG applications, there could be multiple user experience levels, depending on the network transmission quality, etc. Therefore, evaluating a single combination of (PSR, PDB) is not enough since RAN1 does not clearly know its physical meaning and the user experience level it corresponds to.  Proposal 3: RAN1 evaluates multiple combinations of (PSR, PDB) to reflect multiple user experience levels, so that the SI’s outcome is close to real applications and more informative.  Observation 2: If there is no principle/guideline on choosing (PSR, PDB) values, there could be too many combinations and face the following issues:   * Too many combinations will result in large simulation workload * If companies choose quite different values to evaluate, their results are not comparable * The physical meaning and user experience level of each (PSR, PDB) combination is still unclear   Proposal 4: RAN1 discusses and agrees on multiple typical combinations of (PSR, PDB), wherein each combination represents one user experience level. Such combinations of (PSR, PDB) are prioritized in RAN1 evaluations.  Proposal 5: RAN1 agrees on the following table for evaluating multiple combinations of (PSR, PDB):   * The detailed values of packet success rate and PDB of each level will be separately discussed  |  |  |  | | --- | --- | --- | | XR Quality Index (XQI) | Description | (Packet success rate X%, PDB (ms)) | | 5 | Excellent | (X1, T1) | | 4 | Good | (X2, T2) | | 3 | Fair | (X3, T3) | | 2 | Poor | (X4, T4) | | 1 | Bad | (X5, T5) |   Proposal 6: For VR/AR DL video, RAN1 agrees on the following Table 3 for evaluating multiple combinations of (PSR, PDB);  Table 3. XQI table for VR/AR DL video   |  |  |  |  | | --- | --- | --- | --- | | XR Quality Index (XQI) | Description | (Packet success rate X%, PDB (ms)) | | | Single-stream | Multi-stream  {I-stream, P-stream} | | 5 | Excellent | (99, 7) | { (99.5, 7), (95, 7) } | | 4 | Good | (99, 10) | { (99.5, 10), (95, 10) } | | 3 | Fair | (95, 13) | { (95.5, 13), (90, 13) } | | 2 | Poor | (95, 20) | { (95.5, 20), (90, 20) } | | 1 | Bad | (X <95, or PDB>20) | { (X <95.5, or PDB>20),  or (X <90, or PDB>20) } |   Proposal 7: For CG DL video, RAN1 agrees on the following Table 5 for evaluating multiple combinations of (PSR, PDB):  Table 4. XQI table for CG DL video   |  |  |  |  | | --- | --- | --- | --- | | XR Quality Index (XQI) | Description | (Packet success rate X%, PDB (ms)) | | | Single-stream | Multi-stream  {I-stream, P-stream} | | 5 | Excellent | (99, 12) | { (99.5, 12), (95, 12) } | | 4 | Good | (99, 15) | { (99.5, 15), (95, 15) } | | 3 | Fair | (95, 18) | { (95.5, 18), (90, 18) } | | 2 | Poor | (95, 25) | { (95.5, 25), (90, 25) } | | 1 | Bad | (X <95, or PDB>25) | { (X <95.5, or PDB>25), or  (X <90, or PDB>25) } | |

1. **Please share your comments on additional per UE KPI, XR Quality Index (XQI) proposed by Huawei.**

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

## DL Traffic Model: Multiple Streams

1. **DL multiple streams per UE.**

RAN1#104-e agreement

|  |
| --- |
| * FFS if there are multiple streams (if adopted) * FFS the following in RAN1#104-bis-e   + Whether/how to model and evaluate I-frame and P-frame for both DL and UL, e.g., separate definition of fps, packet size, QoS requirements (e.g., PER, PDB), etc.   + Whether/how to separately model and evaluate two streams of video and audio/data for both DL and UL   + Whether/how to model and evaluate FOV (high-resolution) and non-FOV (lower-resolution omnidirectional) streams, e.g., separate definition of fps, packet size, QoS requirements (e.g., PER, PDB), etc |

Companies’ views in RAN1#104bis-e tdocs are presented in the table below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Huawei | Proposal 4: For DL video of AR/VR/CG, adopt M1=2 for modelling I-frame and P-frame separately, and adopt the multi-stream traffic model in following Table 6.  Table 6. Multi-stream model for DL video   |  |  |  | | --- | --- | --- | | 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 | | Traffic arrival 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 | * Stream #1: 1 * Stream #2: N-1 * N is the number of slice per frame, e.g. N = 8. | * Stream #1: 1 * Stream #2: 1 | | Average data rate | Stream #1: Stream #2 = : (N-1) | Stream #1: Stream #2 = : (K-1) | | is average size ratio between one I-frame/slice and one P-frame/slice, e.g. = 2.   * Other values can be optionally evaluated. | |   Note: the QoS requirement for each stream is separately discussed in the KPI part |
| OPPO | Not support to model multiple streams |
| vivo | Proposal 5: For multiple streams XR traffic model in DL, GOP-based/slice-based multiple streams traffic model in Table 2/Table 3 can be considered.  Table 2. GOP-based multiple streams traffic model (FPS=60)   |  |  |  |  | | --- | --- | --- | --- | | **Traffic model** | I-frame | P-frame | Note | | **Date rate (Mbps)** |  |  | GOP length = 1 second | | **Packet size distribution** | Truncated Gaussian distribution | |  | | **Mean packet size (Bytes)** | 272177 | 90725 | The average ratio of I-frame size and P-frame size is around 3:1 | | **STD of packet sizes (Bytes)** | 40826 | 13608 | 15% of Mean packet size | | **Maximum packet size (Bytes)** | 408265 | 136087 | 1.5 \* Mean packet size | | **Minimum packet size (Bytes)** | 900 | 300 |  | | **Packet arrival interval (ms)** | 1000 |  | 1 I-frame and 59 P-frames in one second | | **Packet delay budget (ms)** | 10 | |  |   Table 3. Slice-based multiple streams traffic model (FPS=60)   |  |  |  |  | | --- | --- | --- | --- | | **Traffic model** | I-frame | P-frame | Note | | **Date rate (Mbps)** |  |  |  | | **Packet size distribution** | Truncated Gaussian distribution | |  | | **Mean packet size (Bytes)** | 28125 | 65625 | The average ratio of I-slice size and P-slice size is around 3:1, and each encoded video frame contains 1 I-slice and 7 P-slices | | **STD of packet sizes (Bytes)** | 4218 | 9843 | 15% of Mean packet size | | **Maximum packet size (Bytes)** | 42187 | 98437 | 1.5 \* Mean packet size | | **Minimum packet size (Bytes)** | 93 | 217 |  | | **Packet arrival interval (ms)** |  |  | each encoded video frame contains 1 I-slice and 7 P-slices | | **Packet delay budget (ms)** | 10 | |  |   Proposal 2: For a given data rate, single stream with two-eye buffers can be modelled as:   * Model 1: each packet representing both eyes buffers arrives at the same time at X FPS and the sum of packet size for both eyes is equal to the size of a packet in simulation. * Model 2: packet representing left or right eye buffer arrives at 2\*X FPS and the packet size of left or right eye is the size of a packet in simulation. |
| MTK | Adopt the IDR refresh model for both UL/DL videos for RAN1 evaluation considering its low encoding complexity and wide usage in current industry (e.g. Google Stadia, Nvidia Geforce Now)  Proposal 8: 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  Adopt Tp as PDB for P-Frames and Ti as PDB for I-frames with Tp < Ti.  FFS Tp = 8 ms and Ti = 12 ms |
| Nokia | Proposal 5: Consider a single stream in downlink and single stream in uplink for VR1 and VR2 applications as a baseline.  Proposal 6: Consider a signle stream in downlink and a single stream in uplink for CG application as a baseline.  Proposal 8: Following SA4 input, consider no differentiation between the types of packets/frames as well as FOV/non-FOV as the baseline evaluation of XR/CG applications. |
| Ericsson | [Proposal 5 RAN1 should not model and evaluate I-frame and P-frame separately which will require introducing new traffic parameters.](#_Toc68631143)  [Proposal 6 RAN1 should avoid including multiple streams caused by a frame type, voice traffic, and non-FoV which will increase traffic modelling complexity and evaluation options.](#_Toc68631144) |
| Apple | Two streams (scene/video + audio/data) for downlink for evaluation on AR2. The audio/data flow is modeled as:   * Periodic:   + 10 milliseconds for framing (SA4 input: 10 ms for data stream and 20 ms for audio) * Data rate   + 0.756 Mbps/s or 1.12 Mbps (SA4 input: 256/512 Kbps for audio, 0.5 Mbps for data) * Packet size: constant packet size calculated from periodicity and data rate   End-to-end (mouth-to-ear) latency: 100 ms (SA4 input: 100 ms for both data and audio stream), air interface latency: 30 ms |
| Xiaomi | In XR evaluation, do not differentiate I-frame stream and P-frame stream in DL video stream. |
| Samsung | If the P-frame needs to have smaller PER or PDB that the I-frame, separate models can be defined; otherwise, a single model for the I-frame is used. |
| LGE | Multi-stream is not further considered in DL for VR1/VR2/CR/AR1/AR2 applications |
| AT&T | In addition to M1=1 and M2=1 streams, support M1=2 and M2=2, where in both the DL and UL a user has one video stream based on a Truncated Gaussian packet size distribution and one data/control stream based on a fixed packet size and inter-arrival time |

**Summary**:

* Five companies propose not to evaluate multiple streams per UE in DL.
* Five companies are open to evaluate multiple streams

1. **Based on the discussions and proposals in tdocs for RAN1#104bis-e, the moderator makes the following proposal. Please share your comments.**

**Moderator proposal:**

* In addition to single stream per UE in DL which is baseline, two streams are optionally evaluated.
* FFS details of traffic model, KPI per stream, and per UE KPI.

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

## UL CG/VR

1. **WA for UL CG/VR**

RAN1#104-e agreement

|  |
| --- |
| Working assumption: 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 (SA4 input)   + PDB: 10 ms |

Companies’ views in RAN1#104bis-e tdocs are presented in the table below.

|  |  |
| --- | --- |
| Huawei | Confirm the WA. |
| OPPO | Confirm the WA. |
| vivo | Confirm the WA. |
| Xiaomi | Confirm the WA. |
| Qualcomm | Confirm the WA. |
| ZTE | Confirm the WA. |
| DOCOMO | Confirm the WA. |

**Summary**:

* Seven companies propose to confirm the above WA.

1. **Based on the discussions and proposals in tdocs for RAN1#104bis-e, the moderator makes the following proposal. Please share your comments.**

**Moderator proposal:** 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 (SA4 input)
    - PDB: 10 ms.
* 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 values can be optionally evaluated

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

## UL: AR

1. **UL AR traffic model**

SA4 has discussed the multiple streams for AR applications. As shown in the following table, there are a variety of services in uplink transmission, such as pose, video, audio, data, etc., where video, camera scene, audio and data streams have similar E2E latency requirement. The required bit rate of audio and data is less than the video stream.

Uplink multiple streams for AR application

|  |  |  |
| --- | --- | --- |
| **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)** | **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 RAN1#104bis-e tdocs on UL AR traffic model are presented in the table below.

|  |  |
| --- | --- |
| Huawei | Proposal 6: There are M2=3 streams in UL traffic model of AR, where   * One stream for UL pose/control, * The model for UL pose/control of VR/CG can be reused. * Two streams for UL video to model I-stream and P-stream separately, * The multi-stream model for DL video of VR/AR/CG can be reused. |
| OPPO | One stream for video and another stream for audio or control/pose   * Pose/Control (Same as CG/VR)   + Periodic: 4ms (no jitter)   + Fixed: 100 bytes (SA4 input)   + PDB: 10 ms * UL Video stream   + 10Mbps @60fps (baseline)   + 20Mbps@60fps (optional)   + For packet size, reuse the truncated Gaussian distribution for DL packet size except the mean value is adjusted according to the data rates of UL video. Other parameters are kept the same   + For jitter, reuse the truncated Gaussian distribution for DL packet arrival jitter. All parameters are kept the same   + 10ms (baseline)     - Other values can be evaluated optionally   A UE is declared as satisfied only when all streams meets their corresponding requirements  Not support to model and evaluate I-frame and P-frame for the evaluation of XR/CG on NR |
| vivo | * Option 1: single pose stream. * Option 2: single video stream. * Option 3: two streams with pose/control and video streams   Table 5. Single stream traffic model of video in UL   |  |  |  | | --- | --- | --- | | **Traffic model** | Video | Note | | **Data rate (Mbps)** | 20 |  | | **Frame per second** | 60 |  | | **Packet size distribution** | Truncated Gaussian distribution |  | | **Mean packet size (Bytes)** | 41667 | Average data rate / FPS / 8 [Bytes] | | **STD of packet sizes (Bytes)** | 6250 | 15% of Mean packet size | | **Maximum packet size (Bytes)** | 62500 | 1.5 \* Mean packet size | | **Minimum packet size (Bytes)** | 4167 | 0.1 \* Mean packet size | | **Packet arrival interval (ms)** | 16.67 |  | | **Packet delay budget (ms)** | 60 | 60ms for AR | |
| MTK | * M2=2 or 3 for AR to model video and control/pose separately   No need to model the audio stream separately |
| Futurewei | * A single video stream for a UE: periodic with 60 fps, no jitter * Average data rate: 20 Mbps @ 60 fps (baseline) * Truncated Gaussian distribution is used for the packet size distribution of video stream for UL AR with the following parameters:   + Mean: derived from fps and average data rate   + STD: 15% of Mean packet size   + Max packet size: 1.5 x Mean packet size   + Min packet size: 0.5 x Mean packet size * PDB: 60 ms (baseline) |
| Nokia | Adopt a single stream of video in UL for AR2: XR Conversational as a baseline. The average data rate is 10 Mbit/s (1080p) and the frame rate is 60 fps. The PDB is 10 ms.  Proposal 4: No jitter is assumed for the UL video stream.  Proposal 7: Consider a single stream in downlink and a single stream in uplink for AR application as a baseline. Any additional streams consider as optional. |
| Ericsson | [Proposal 1 The bit rates for AR UL scene can be the range of 2Mbps to 20Mbps and the latency requirement is similar as DL AR/VR video, i.e., 5ms to 20ms.](#_Toc68631139)  [Proposal 5 RAN1 should not model and evaluate I-frame and P-frame separately which will require introducing new traffic parameters.](#_Toc68631143)  [Proposal 6 RAN1 should avoid including multiple streams caused by a frame type, voice traffic, and non-FoV which will increase traffic modelling complexity and evaluation options.](#_Toc68631144) |
| Xiaomi | An UL pose stream and a single UL video data stream are used as UL traffic model for AR2 use case |
| Apple | 3 streams (scene/video + audio/data + pose/control) for UL and two streams (scene/video + audio/data) for downlink can be used for evaluation on AR2. The audio/data flow is modeled as:   * Periodic:   + 10 milliseconds for framing (SA4 input: 10 ms for data stream and 20 ms for audio) * Data rate   + 0.756 Mbps/s or 1.12 Mbps (SA4 input: 256/512 Kbps for audio, 0.5 Mbps for data) * Packet size: constant packet size calculated from periodicity and data rate   End-to-end (mouth-to-ear) latency: 100 ms (SA4 input: 100 ms for both data and audio stream), air interface latency: 30 ms |
| Qualcomm | * Stream 1: pose/control (same as VR/CG) * Stream 2: aggregated stream for scene, video, data, and audio.   + Truncated Gaussian distribution for Packet size with same parameter values |
| Samsung | 2 UL streams (pose and scene upload) |
| Sony | 3 streams for UL (pose and two cameras) |
| LGE | Multi-stream is not further considered in UL for VR1/VR2/CG applications   * In the case where two streams are used for UL traffic modelling, one for control/pose and the other for video stream/scene update/audio/data, a UE is declared a satisfied UE if more than X1 (%) of packets for control/pose are successfully transmitted within a given air interface PDB1, AND more than X2 (%) of packets for video stream/scene update/audio/data are successfully transmitted within a given air interface PDB2.   + X1 is [99] and X2 is [95]   + PDB1 is 10ms and PDB2 is [100]ms   Proposal 5: UL Traffic model for video stream/scene update/audio/data   * Frame per second (fps)   + 60 fps (SA4 input) – no jitter * Average data rate for UL video stream:   + 20 Mbps (~half of the average of VR/AR average data rate for DL video stream) * Truncated Gaussian distribution is used for the packet size distribution   + Parameters of Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)     - Mean: Derived from average data rate and fps as follows.       * (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]     - STD: same as in DL     - Max packet size: same as in DL     - Min packet size: FFS (need input from SA4) * PDB   + [100] ms (based on SA4 input) |
| InterDigital | 2 different traffic flows in UL as mandatory for AR evaluations. FFS for using more than 2 dfferent traffic flows in UL for AR |
| AT&T | support M2=2, where a user has one video stream based on a Truncated Gaussian packet size distribution and one data/control stream based on a fixed packet size and inter-arrival time |
| DOCOMO | * *Adopt two streams for UL for AR applications*   + *Traffic model for pose/control information can be same as CG/VR.*   + *Traffic model for scene update/video/audio data,*     - *Periodicity: 60 fps*     - *Data rate: 20 Mbps*     - *PDB: 60 ms* |

**Summary**:

* Ten companies support evaluation of two streams, one for pose/control and the other for scene update/video/audio/data
* Four companies propose to evaluate only one stream for UL AR.
* One company propose 3 streams (scene/video + audio/data + pose/control).
* One company propose 3 streams for UL (pose and two cameras)

1. **Based on the discussions and proposals in tdocs for RAN1#104bis-e, the moderator makes the following proposal. Please share your comments.**

**Moderator proposal:**

* For UL AR,
  + Baseline: two streams are defined as follows.
    - Stream 1: pose/control
      * Traffic model and QoS parameters are same as those for pose/control for UL CG/VR.
    - Stream 2: aggregated stream for scene, video, data, and audio.
      * Packet size: Truncated Gaussian distribution with the same parameter values with DL (Mean, STD, Max, Min)
      * Periodicity: 60 fps
      * Data rate: 10 Mbps (baseline), 20 Mbps (optional)
      * PDB: 60 ms
  + Single stream (Stream 2 above) and/or more than two streams can be optionally evaluated.
  + A UE is declared as satisfied only when each stream meets the following requirement.
    - More than X (%) of packets are successfully delivered within the given air interface PDB
      * The baseline X value is 99.
      * Other values can be optionally evaluated

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

## Others

1. **Other issues**
2. **Please discuss other issues that are not discussed above.**

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

# Summary

# List of contributions in RAN1 #104b-e

1. [R1-2102320](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102320.zip) Traffic model for XR and Cloud Gaming Huawei, HiSilicon
2. [R1-2102418](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102418.zip) Discussion on the XR traffic models for evaluation OPPO
3. [R1-2102546](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102546.zip) Discussion on traffic models of XR vivo
4. [R1-2102616](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102616.zip) XR traffic model CATT
5. [R1-2102686](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102686.zip) Traffic Model for XR and CG MediaTek Inc.
6. [R1-2102769](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102769.zip) XR traffic model FUTUREWEI
7. [R1-2102827](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102827.zip) On Traffic Model for XR study Nokia, Nokia Shanghai Bell
8. [R1-2102955](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102955.zip) Traffic model for XR Ericsson
9. [R1-2102969](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2102969.zip) Discussion on Traffic Model for XR services Xiaomi
10. [R1-2103054](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103054.zip) Traffic Model for XR Intel Corporation
11. [R1-2103128](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103128.zip) Views on XR traffic model Apple
12. [R1-2103192](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103192.zip) Remaining Issues on XR Traffic Models Qualcomm Incorporated
13. [R1-2103264](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103264.zip) Traffic model for XR Samsung
14. [R1-2103278](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103278.zip) Further Discussion on Traffic Model for XR Evaluations ZTE, Sanechips
15. [R1-2103317](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103317.zip) Considerations on XR traffic model Sony
16. [R1-2103360](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103360.zip) Discussion on traffic models for XR evaluation LG Electronics
17. [R1-2103429](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103429.zip) UL traffic flows for XR applications InterDigital, Inc.
18. [R1-2103437](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103437.zip) XR Traffic Model Considerations AT&T
19. [R1-2103598](file:///C:\Users\wanshic\OneDrive%20-%20Qualcomm\Documents\Standards\3GPP%20Standards\Meeting%20Documents\TSGR1_104b\Docs\R1-2103598.zip) Discussion on traffic model for XR NTT DOCOMO, INC.

# Appendix-A (proposals in RAN1#104bis-e tdocs)

**Huawei, HiSilicon**

*Observation 1: For a given VR video, the parameters of the video packet size distribution are related to video encoding configurations, e.g. error resilience, rate control, etc.*

*Observation 2: In the frame-based I/P-stream model for AR/VR/CG, the packet arrival of I-stream and P-stream has a Group-Of-Pictures (GOP) structure.*

*Observation 3: In the slice-based I/P-stream model for AR/VR/CG, both streams have periodic traffic with packet arrival interval 1/FPS.*

*Proposal 1: The following parameters for truncated Gaussian distribution for packet size can be a starting point (note: these parameter values are those before the truncation)*

* *STD: 15% of Mean packet size*
* *Max packet size: 2 \* Mean packet size*
* *Min packet size : 25% of Mean packet size*

*Proposal 2: The following parameters for mean, STD and range of jitter for DL video stream can be a starting point for initial evaluation.*

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

*Proposal 3: If jitter is considered, the remaining scheduling time of a packet is affected by jitter, i.e. remaining scheduling time = air interface PDB – jitter.*

*Proposal 4: For DL video of AR/VR/CG, adopt M1=2 for modelling I-frame and P-frame separately, and adopt the multi-stream traffic model in following Table 6.*

*Table 6. Multi-stream model for DL video*

|  |  |  |
| --- | --- | --- |
| *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* |
| *Traffic arrival 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* | * *Stream #1: 1* * *Stream #2: N-1* * *N is the number of slice per frame, e.g. N = 8.* | * *Stream #1: 1* * *Stream #2: 1* |
| *Average data rate* | *Stream #1: Stream #2 = : (N-1)* | *Stream #1: Stream #2 = : (K-1)* |
| *is average size ratio between one I-frame/slice and one P-frame/slice, e.g. = 2.*   * *Other values can be optionally evaluated.* | |

*Note: the QoS requirement for each stream is separately discussed in the KPI part.*

*Proposal 5: Confirm the following working assumption on traffic model for UL pose/control of CG/VR.*

* *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 (SA4 input)*
    - *PDB: 10 ms*

*Proposal 6: There are M2=3 streams in UL traffic model of AR, where*

* *One stream for UL pose/control,*
* *The model for UL pose/control of VR/CG can be reused.*
* *Two streams for UL video to model I-stream and P-stream separately,*
* *The multi-stream model for DL video of VR/AR/CG can be reused.*

**OPPO**

Proposal 1: For the evaluation of XR/CG on NR, the case where M1=1 and M1=2 stream in DL and UL is mandatory

* + Not support to model multiple streams
  + If RAN1 agrees to support M1>1 or M2>1 for evaluation, companies can also evaluate M1=2 and/or M2=2 optionally
    - One stream for video
    - Another stream for audio or control/pose

Proposal 2: If multiple streams are used in DL and/or UL, a UE is declared as satisfied only when all streams meets their corresponding requirements.

Proposal 3: Not support to model and evaluate I-frame and P-frame for the evaluation of XR/CG on NR.

Proposal 4: For the case of 120fps, companies can select one or more values for evaluation

* VR/AR: 60, 90, 120 Mbps
* CG: 16, 60, 90 Mbps

Proposal 5: For the distribution of packet sizes, confirm the working assumption of truncated Gaussian distribution by removing the brackets, i.e.,

* STD: 15% or 20% of Mean packet size
* Max packet size: 1.5 x Mean packet size
* Min packet size: 0.5 x Mean packet size
  + Companies can report whether this parameter is used or not

Proposal 6: For the distribution of inter-packet arrival jitter, adopt the truncated Gaussian distribution with the following parameters:

* Mean: 0
* STD: 3ms
* Range: (-5ms, 5ms)

Proposal 7: For air interface PDB for DL video stream, no more mandatory value is needed.

Proposal 8: For the UL traffic of AR conversational, support the following two different models:

* Pose/Control (Same as CG/VR)
  + Periodic: 4ms (no jitter)
  + Fixed: 100 bytes (SA4 input)
  + PDB: 10 ms
* Data rate for UL Video stream
  + 10Mbps @60fps (baseline)
  + 20Mbps@60fps (optional)

Proposal 9: For UL video stream of AR conversational:

* For packet size, reuse the truncated Gaussian distribution for DL packet size except the mean value is adjusted according to the data rates of UL video. Other parameters are kept the same
* For jitter, reuse the truncated Gaussian distribution for DL packet arrival jitter. All parameters are kept the same

Proposal 10: For UL video stream of AR conversational, the air interface PDB

* 10ms (baseline)
* Other values can be evaluated optionally

**vivo**

Observation 1: For multiple streams modelling, it is necessary to consider I-frame stream and P-frame stream in XR traffic model.

Observation 2: There is no need to model audio stream separately in XR traffic model.

Observation 3: FOV and non-FOV streams based XR traffic model are similar to I-frame and P-frame streams based XR traffic model.

Proposal 1: For the association between jitter and PDB, actual PDB = (ideal PDB – jitter) for each packet.

Proposal 2: For a given data rate, single stream with two-eye buffers can be modelled as:

* Model 1: each packet representing both eyes buffers arrives at the same time at X FPS and the sum of packet size for both eyes is equal to the size of a packet in simulation.
* Model 2: packet representing left or right eye buffer arrives at 2\*X FPS and the packet size of left or right eye is the size of a packet in simulation.

Proposal 3: For XR traffic model in DL, the two traffic models in Table 1 are considered as the starting point for XR evaluation.

Proposal 4: Confirm the working assumptions on the truncated Gaussian distribution for packet size and jitter modelling.

* Parameters of Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)
  + Mean: Derived from average data rate and fps as follows.
    - (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
  + STD
    - 15% of Mean packet size derived above
  + Max packet size
    - 1.5 x Mean packet size derived above
  + Min packet size
    - 0.1 \* Mean packet size derived above
* Jitter for DL video stream for a single UE
  + Per the agreed statistical traffic model, arrival time of packet k is k/X \* 1000 [ms] + J [ms], where X is the given fps value and J is a random variable.
  + J is drawn from a truncated Gaussian distribution:
    - Mean: 0
    - STD: 2 ms
    - Range: [-4, 4]ms

Proposal 5: For multiple streams XR traffic model in DL, GOP-based/slice-based multiple streams traffic model in Table 2/Table 3 can be considered.

Proposal 6: Confirm the working assumption of UL single stream traffic model for pose/control.

Proposal 7: For UL single stream traffic model for video, the traffic model in Table 5 is supported at least for AR.

Proposal 8: UL multiple streams with both pose/control and video streams are supported for UE power consumption evaluation.

Proposal 9: For XR DL traffic model, consider the following options:

* Option 1: single video stream.
* Option 2: two streams with I-frame and P-frame.

Proposal 10: For XR UL traffic model, consider the following options:

* Option 1: single pose stream.
* Option 2: single video stream.
* Option 3: two streams with pose/control and video streams.

Proposal 11: A UE with multiple streams is declared as a satisfied UE if each stream from the multiple streams 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.

Proposal 12: When two streams are modeled for a user in UL, the stream for pose/control information may have an X value of 99.9 and a given air interface PDB of 10ms, and the stream for scene information may have an X value of 99 and a given air interface PDB of 60ms.

Proposal 13: When the DL video traffic is divided into two streams, e.g. one stream for I-frames and the other for P-frames, the stream for I-frames may have an X value of 99, and the stream for P-frames may have an X value of 95, but the two streams may share the same given air interface PDB of 10ms.

**CATT**

Observation 1: The truncated Gaussian distribution can be used for modelling the packet size for XR and parameters are those of Gaussian distribution before truncation.

Observation 2: It observes that

* + The average data rates based on the mean values are 13.68~14.19 Mbps for VR2.
  + The STD of Packet size is 3.32%~8.33% for VR2, which would be no more than 10% of Mean packet size.
  + The maximum packet sizes are limited by the transmission characteristics, such as “30Mbit/s capped VBR with window 200ms” and “30Mbit/s CBR with window 1 frame”, which is about 1.06~1.14 x Mean packet size.
  + The minimum packet size would be at least larger than the minimum IP packet size, i.e. 46Bytes.

Observation 3: There are two alternatives for modeling the jitter from the different aspects:

* + Opt1-Frame Delay (J): The arrival time of packet k is k/X×1000 [ms] + J [ms], where X is the given FPS value and J is a random variable.
  + Opt2-Inter Arrival Time Jitter (JJ): The inter arrival time between the packet k and the packet k+1 is 1/X×1000 [ms] + JJ [ms], where X is the given FPS value and JJ is a random variable.

Observation 4: For Opt1-Frame Delay (J), the absolute arrival time of packet k is k/X×1000 [ms] + J [ms], where X is the given FPS value and J is a random variable. It observes that

* + The statistic distribution for Opt1-Frame Delay (J) is close to the uniform distribution.
  + The value of Opt1-Frame Delay (J) would be always positive.
  + The mean value is 19.8 ms, which is not equal to 0.
  + The STD is 5.71-5.84 ms.
  + The range of Opt1-Frame Delay (J) is [9.09, 30.68] ms.

Observation 5: For Opt2-Inter Arrival Time Jitter (JJ), the inter arrival time between the packet k and the packet k+1 is 1/X×1000 [ms] + JJ [ms], where X is the given FPS value and JJ is a random variable. It observes that

* + Opt2-Inter Arrival Time Jitter (JJ) is characterized by the truncated Gaussian distribution.
  + The value of Opt2-Inter Arrival Time Jitter (JJ) could be either positive or negative.
  + The mean value is 0 ms.
  + The STD is 8.10~8.19 ms.
  + The range of Opt2-Inter Arrival Time Jitter (JJ) is [-19.54, 19.87] ms.
  + The percentage of packet arrival out of order, i.e. the Opt2-Inter Arrival Time Jitter (JJ) less than -1/X×1000 [ms], is 1.06%~1.56%.

Proposal 1:

* + The truncated Gaussian distribution should be used for modelling the packet size of video stream with VBR and the uniform distribution should be used for modelling the packet size of video stream with CBR.

Proposal 2: For the parameters of the statistical distribution for Packet size

* + The mean value of packet size could be derived by the statistical traffic modelling that (average data rate) / (FPS for video stream, i.e., # packets per second in our statistical model) / 8 [bytes].
  + The STD of Packet size could be derived by 5% and 10% of Mean packet size for the video stream with CBR and VBR, respectively.
  + The maximum packet size could be 1.2 time of and equal to Mean packet size for the video stream with VBR and CBR, respectively.
  + The minimum packet size could be limited by the minimum IP packet size, i.e. 46Bytes.

Proposal 3: Either two alternatives can be used for jitter modeling.

* + Opt1-Frame Delay (J): The arrival time of packet k is k/X×1000 [ms] + J [ms], where X is the given FPS value and J is a random variable.
  + Opt2-Inter Arrival Time Jitter (JJ): The inter arrival time between the packet k and the packet k+1 is 1/X×1000 [ms] + JJ [ms], where X is the given FPS value and JJ is a random variable.

Proposal 4: If jitter is modeled as Opt1-Frame Delay (J), in which the arrival time of packet k is k/X×1000 [ms] + J [ms] under the given FPS value, the following parameters could be considered.

* + The uniform distribution is used for modelling the random variable J.
  + Mean: 20 ms
  + STD: 6.35 ms
  + Range: [9, 31] ms

Proposal 5: If jitter is modeled as Opt2-Inter Arrival Time Jitter (JJ), in which the inter arrival time between the packet k and the packet k+1 is 1/X×1000 [ms] + JJ [ms] under the given FPS value, the following parameters could be considered.

* + The truncated Gaussian distribution is used for modelling the random variable JJ.
  + Mean: 0
  + STD: 8 ms
  + Range: [-1/X×1000, 20] ms

**MediaTek Inc.**

Observation 1: CG and XR display different traffic types within the same application, in both UL and DL directions

Observation 2: 5GS system awareness of differentiated frames may be beneficial

Proposal 1: Adopt the IDR refresh model for both UL/DL videos for RAN1 evaluation.

Proposal 2: traffic model shall take into account different traffic types and possibly differentiated frames within the same application, in both UL and DL directions

Proposal 3: In terms of the values of M1 & M2 for evaluation of DL/UL

* For DL:
  + M1=2 to model I-frame and P-frame separately with different QoS requirements for VR/AR/CG
* For UL:
  + M2=1 for VR/CG (agreed in RAN1 #104e)
  + M2=2 or 3 for AR to model video and control/pose separately

Proposal 4: No need to model the audio stream separately

Proposal 5: Adopt the same PER requirements for I-frames and P-frames.

* FFS 99%

Proposal 6: Adopt Tp as PDB for P-Frames and Ti as PDB for I-frames with Tp < Ti.

* FFS Tp = 8 ms and Ti = 12 ms.

Proposal 7: Coordinate and cooperate with SA4 to construct a video quality evaluation block to use in both RAN1 and SA4 to evaluate the proposed QoS requirements and the proposed enhancement.

Proposal 8: 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.

Proposal 9: Confirm Jitter working assumptions.

* + J is drawn from a truncated Gaussian distribution:
    - Mean: 0
    - STD: 2 ms
    - Range: [-4, 4]ms
      * Note: The values ensure that packet arrivals are in order (i.e., arrival time of a next packet is always larger than that of the previous packet)

Proposal 10: Distinguish Jitter parameters depending on XR/CG server location (Edge, Cloud)

Proposal 11: Confirm Packet size working assumptions.

* + Packet Size is drawn from a truncated Gaussian distribution:
    - Mean: (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
    - STD: 15% of Mean packet size derived above
    - Range: 1.5 × Mean packet size derived above

**FUTUREWEI**

Proposal 1: FS\_NR\_XR\_eval adopts the following regarding the parameters of truncated Gaussian distribution for packet size:

* STD: 15% of Mean packet size
* Max packet size: 1.5 x Mean packet size
* Min packet size: 0.5 x Mean packet size

Proposal 2: FS\_NR\_XR\_eval supports modeling single stream/flow on each direction as a baseline.

Proposal 3: FS\_NR\_XR\_eval adopts the following regarding the UL traffic model for AR:

* A single video stream for a UE: periodic with 60 fps, no jitter
* Average data rate: 20 Mbps @ 60 fps (baseline)
* Truncated Gaussian distribution is used for the packet size distribution of video stream for UL AR with the following parameters:
  + Mean: derived from fps and average data rate
  + STD: 15% of Mean packet size
  + Max packet size: 1.5 x Mean packet size
  + Min packet size: 0.5 x Mean packet size
* PDB: 60 ms (baseline)

**Nokia, Nokia Shanghai Bell**

Proposal 1: Adopt the following parameters for the packet (frame) size distribution:

* STD: 2% of mean packet (frame) size
* Max packet (frame) size: 1.1 x mean packet (frame) size
* Min packet (frame) size: 0.9 x mean packet (frame) size

Other values for the packet (frame) size distribution are optional.

Proposal 2: Adopt the following parameters for jitter:

* Mean: 0
* STD: 3 ms
* Range: (-6, 6) ms

Other values for the jitter distribution are optional.

Proposal 3: Adopt a single stream of video in UL for AR2: XR Conversational as a baseline. The average data rate is 10 Mbit/s (1080p) and the frame rate is 60 fps. The PDB is 10 ms.

Proposal 4: No jitter is assumed for the UL video stream.

Proposal 5: Consider a single stream in downlink and single stream in uplink for VR1 and VR2 applications as a baseline.

Proposal 6: Consider a signle stream in downlink and a single stream in uplink for CG application as a baseline.

Proposal 7: Consider a single stream in downlink and a single stream in uplink for AR application as a baseline. Any additional streams consider as optional.

Proposal 8: Following SA4 input, consider no differentiation between the types of packets/frames as well as FOV/non-FOV as the baseline evaluation of XR/CG applications

**Ericsson**

[Observation 1 The bit rates requirement of AR UL scene can be lower than VR/AR DL video while the latency requirement of it is similar as VR/AR DL video.](#_Toc68631137)

[Observation 2 Differentiating and evaluating I-frame and P-frame separately is not essential from a XR traffic characteristics and requirement perspective.](#_Toc68631138)

Based on the discussion in the previous sections we propose the following:

[Proposal 1 The bit rates for AR UL scene can be the range of 2Mbps to 20Mbps and the latency requirement is similar as DL AR/VR video, i.e., 5ms to 20ms.](#_Toc68631139)

[Proposal 2 The min value of packet size should be the 50% of mean packet size in order to make a symmetric distribution. STD and the max value of packet size are 15% and 150% of mean packet size, respectively.](#_Toc68631140)

[Proposal 3 Confirm the proposed values for Mean, STD, range of the jitter distribution.](#_Toc68631141)

[Proposal 4 A UE is satisfied if more than 99% of packets are successfully received within a given air interface PDB.](#_Toc68631142)

[Proposal 5 RAN1 should not model and evaluate I-frame and P-frame separately which will require introducing new traffic parameters.](#_Toc68631143)

[Proposal 6 RAN1 should avoid including multiple streams caused by a frame type, voice traffic, and non-FoV which will increase traffic modelling complexity and evaluation options.](#_Toc68631144)

**Xiaomi**

Proposal 1: Audio/Data stream is not considered in DL video stream for VR2, CG and AR2 services.

Proposal 2: In XR evaluation, do not differentiate I-frame stream and P-frame stream in DL video stream.

Proposal 3: For XR DL evaluation, a single DL video stream including frames for both eye buffers is assumed.

- Interleaved eye buffer model can be optionally considered.

Proposal 4: Send LS to SA4 to confirm on the working assumption of packet size & jitter distribution.

Proposal 5: The initial frame generation time should be randomized among different UEs.

Proposal 6: Confirm the working assumption on UL traffic model and QoS parameters for CG/VR and Pose/control.

Proposal 7: An UL pose stream and a single UL video data stream are used as UL traffic model for AR2 use case.

Proposal 8: For per UE KPI, the exact value of X is set to be 99.9

**Intel Corporation**

Observations-1:

* for CBR configurations the frame-size variations are quite small
  + the max/mean frame-size ratio is ~ 1.06
  + the min/mean frame-size ratio is ~0.93
  + the std/mean frame-size ratio is ~0.02
* for cVBR configurations the frame-size variations are larger
  + the max/mean frame-size ratio is ~ 1.18 – 1.94 with smaller ratios corresponding to 8 slice/eye buffer case while large ratio corresponding to 1 slice case
  + the min/mean frame-size ratio is ~ 0.24 – 0.48 with larger ratios corresponding to 8 slice/eye buffer case while small ratio corresponding to 1 slice case
  + the std/mean frame-size ratio is ~ 0.07 – 0.14

Observations-2:

We observe that

* there is significant variation in traffic observed at L2 (P-trace) for a given content model (V-trace) and use-case example: VR2 30 Mbps 2 eye buffers at 2Kx2K at 60 fps, 8bit
* this variation is due to encoding model (slices, encoding delay, etc.) and content delivery model (packetization etc.)
* distribution of frame-size based on P-trace is asymmetric (heavy tail below mean, light tail above mean)

Observations-3:

Comparing V-trace and P-trace frame-sizes we can observe that:

* distribution of frame-size based on P-trace is asymmetric (heavy tail below mean, light tail above mean) while frame-size based on V-trace is more symmetric in shape
* the max/mean frame-size calculated based on V-trace is much larger than that calculated based on P-trace. The encoding and the content delivery model clearly affects the frame-size distribution.

Observations-4:

The current frame-based statistical model lacks the following:

* different distributions for IP packet sizes (limited, unlimited, etc.)
* variation in the number of packets per burst, even for a given frame-size
* variation in burst length, even for a given frame-size
* assymmetry in frame-size distribution below and above mean
* variation due to use-cases (CG video is envisioned to be more interactive than VR)

Proposal-1: Consider more accurate trace-based traffic model leveraging the SA4 work in RAN1 XR simulations

**Apple**

Observation 1: From SA4 traffic model on XR conversational, it is clear that uplink traffic is with substantial throughput requirements.

Observation 2: 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.

Proposal 3: 3 streams (scene/video + audio/data + pose/control) for uplink and two streams (scene/video + audio/data) for downlink can be used for evaluation on AR2. The audio/data flow is modeled as:

* Periodic:
  + 10 milliseconds for framing (SA4 input: 10 ms for data stream and 20 ms for audio)
* Data rate
  + 0.756 Mbps/s or 1.12 Mbps (SA4 input: 256/512 Kbps for audio, 0.5 Mbps for data)
* Packet size: constant packet size calculated from periodicity and data rate

End-to-end (mouth-to-ear) latency: 100 ms (SA4 input: 100 ms for both data and audio stream), air interface latency: 30 ms

**Qualcomm Incorporated**

**Proposal 1**: Adopt the following for DL video streaming

* Parameters of Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)
  + Mean: Derived from average data rate and fps as follows.
    - (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
  + STD
    - 7.5% of Mean
  + Max packet size
    - 135% of Mean
  + Min packet size
    - 54.5% of Mean

**Proposal 2**: Confirm the following WA.

* Jitter for DL video stream for a single UE
  + Per the agreed statistical traffic model, arrival time of packet k is k/X x 1000 [ms] + J [ms], where X is the given fps value and J is a random variable.
  + J is drawn from a truncated Gaussian distribution:
    - Mean: 0 ms
    - STD: 2 ms
    - Range: [4, 4] ms
    - Other values can be optionally evaluated

**Proposal 3**: Evaluate two streams in UL for AR as follows.

* Stream 1: pose/control (same as VR/CG)
  + Periodic: 4ms (no jitter)
    - Other values can be optionally evaluated.
  + Fixed: 100 bytes (SA4 input)
  + PDB: 10 ms
* Stream 2: aggregated stream for scene, video, data, and audio.
  + Traffic model is same as DL video stream, i.e.,
    - Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)
      * Mean: Derived from average data rate and fps as follows.
        + (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
      * STD
        + 7.5% of Mean
      * Max packet size
        + 135% of Mean
      * Min packet size
        + 54.5% of Mean

**Proposal 4**: Adopt X = 99 in the following except for the case when I-frames and P-frames are separately evaluated.

* 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, e.g., 99, 95
  + FFS different values for I-frame and P-frame if evaluation of them is agreed.

Other values can be optionally evaluated

**Samsung**

Proposal 1: XR traffic models consider 1 DL stream (video) and 2 UL streams (pose and scene upload).

Proposal 2: Confirm the frame-level modeling for packet arrivals.

Proposal 3: If the P-frame needs to have smaller PER or PDB that the I-frame, separate models can be defined; otherwise, a single model for the I-frame is used.

Proposal 4: Confirm the standard deviation and the maximum packet size for the truncated Gaussian distribution as 15% and 1.5x of the mean packet size, respectively.

Proposal 5: Jitter can be evaluated based on a truncated Gaussian distribution or can be abstracted from evaluations that can be instead without jitter for few PDB values - scaling by a jitter distribution can then apply.

Proposal 6: For KPIs, the percentage of UEs that can achieve a target PER and a target PDB suffices for data packets. For PDCCH-based scheduling, whether and how the PDCCH BLER does not impact the target PDB needs to also be considered

**ZTE, Sanechips**

[Observation 1: With Alt 1, the ratio between standard deviation and mean value is 0.08 under the configuration of VR2-1, VR2-2 and VR2-6, while the ratio is around 0.13 under the configuration of VR2-5, when bit rate is 30Mbps.](#_Toc68641007)

[Observation 2: With Alt 1, the ratio between standard deviation and mean value is 0.15 under the configuration of VR2-7, VR2-8, when bit rate is 45Mbps.](#_Toc68641008)

[Observation 3: With Alt1, the ratio between the maximal value and mean value is 1.24 under the configuration of VR2-1, VR2-2 and VR2-6, while the ratio is around 1.39 under the configuration of VR2-5, when bit rate is 30Mbps.](#_Toc68641009)

[Observation 4: With Alt 1, the ratio between standard deviation and mean value is 1.45 under the configuration of VR2-7, VR2-8, when bit rate is 45Mbps.](#_Toc68641010)

[Observation 5: Non-negligible bias could be observed between the CDF curves of the distribution and that of the data samples in the range of 5%-95%.](#_Toc68641011)

[Observation 6: The ratio between standard deviation and mean value is ranging from 4.14% to 4.66% in Gaussian distribution of single eye packet size.](#_Toc68641012)

[Observation 7: The ratio between standard deviation and mean value is ranging from 2.27% to 3.14% in Gaussian distribution of double eyes packet size.](#_Toc68641013)

[Observation 8: To attain jittering information as defined in the CSV files available from [2] , subtraction could be performed between the time\_stamp\_in\_micro\_s value of a representative, e..g, the last fraction to the corresponding rendering time.](#_Toc68641014)

[Observation 9: The values in the WA do not comply with the numerical evaluations](#_Toc68641015)

[Observation 10: Packet loss information and packet delay information cannot provide additional information.](#_Toc68641016)

[Observation 11: If multiple data streams are adopted for DL traffic, the difference of XR/CG source related information may influence the scheduling/collision handling of the different streams](#_Toc68641017)

[Observation 12: Frame based and slice based intra refreshing have an impact on the traffic model aspects such as jittering modelling and inter frame arrival time.](#_Toc68641018)

[Proposal 1: Standard deviation and maximal packet size for DL video streaming traffic are determined as follows:](#_Toc68618182)

[ Single eye packet size](#_Toc68618183)

[- STD = 4% \* mean, MAX = 112% \* mean](#_Toc68618184)

[ Dual eye packet size](#_Toc68618185)

[- STD = 3% \* mean, MAX = 109% \* mean.](#_Toc68618186)

[Note: Minimum file size is not considered](#_Toc68618187)

[Proposal 2: Further discuss in RAN1 the jittering related information for DL video streaming including mean/variance/maximal value using the statistics as starting point.](#_Toc68618188)

[Table 5 Summary of VR2 Jitter Statistics](#_Toc68618189)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Mean (ms)** | **STD (ms)** | **Range [ms, ms]** |
| **VR2-1** | -1.358 | 8.16 | [-25.0740, 18.2330] |
| **VR2-2** | 4.231 | 7.0279 | [-16.0640, 19.8290] |
| **VR2-5** | -4.149 | 7.6014 | [-29.2700, 21.0710] |
| **VR2-6** | 6.697 | 7.0095 | [-18.0640, 22.1750] |
| **VR2-7** | 1.955 | 6.9989 | [-19.0440, 18.3690] |
| **VR2-8** | 0.0489 | 7.8489 | [-23.0550, 18.1960] |

[Proposal 3: Consider the reliability requirement as 95%, i.e. the baseline for per UE KPI is updated as](#_Toc68618191)

[A UE is declared a satisfied UE if more than 99 (%) of packets are successfully transmitted within a given air interface PDB.](#_Toc68618192)

[Proposal 4: When determining a XR/CG user is satisfied or not, the following factors are not considered.](#_Toc68618193)

[ Packet loss information](#_Toc68618194)

[ Packet delay information](#_Toc68618195)

[Proposal 5: Confirm the WA on UL traffic of 100Byte packet size, 4ms periodicity as well as 100ms PDB](#_Toc68618196)

[Proposal 6: Standard deviation and maximal packet size for UL video streaming traffic are determined as follows:](#_Toc68618197)

[ Single eye packet size](#_Toc68618198)

[- STD = 4% \* mean, MAX = 112% \* mean](#_Toc68618199)

[ Dual eye packet size](#_Toc68618200)

[- STD = 3% \* mean, MAX = 109% \* mean.](#_Toc68618201)

[Note: Minimum file size is not considered](#_Toc68618202)

[Proposal 7: Further discuss in RAN1 the jittering related information for UL video streaming including mean/variance/maximal value using the statistics as starting point.](#_Toc68618203)

[Table 5 Summary of VR2 Jitter Statistics](#_Toc68618204)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Mean (ms)** | **STD (ms)** | **Range [ms, ms]** |
| **VR2-1** | -1.358 | 8.16 | [-25.0740, 18.2330] |
| **VR2-2** | 4.231 | 7.0279 | [-16.0640, 19.8290] |
| **VR2-5** | -4.149 | 7.6014 | [-29.2700, 21.0710] |
| **VR2-6** | 6.697 | 7.0095 | [-18.0640, 22.1750] |
| **VR2-7** | 1.955 | 6.9989 | [-19.0440, 18.3690] |
| **VR2-8** | 0.0489 | 7.8489 | [-23.0550, 18.1960] |

[Proposal 8: It's expected from SA that the 5QI values shall be finalized before RAN1 could start the discussion regarding the differentiation of the multiple streams.](#_Toc68618205)

**Sony**

Proposal 1: Support AR2 application as communicated from SA4 with 5 streams for UL and 3 for DL. If simplications are desired, we could have an option with 3 streams for UL (pose and two cameras) and 1 stream for DL (only video).

Proposal 2: Define air interface delay for uplink that is measured from the point when a packet is transmitted by the UE to the point when it is successfully delivered to gNB. FFS: the reference transmission point at the UE side (e.g. TX antenna connector, etc).

Proposal 3: Air interface PDB UL for VR/AR is 10 ms.

Proposal 4: Use the media characteristics / simulation configurations that SA4 already defined in LS R1-2101765.

Observation 1: Separate modelling of I-frame and P-frame is desirable to provide accurate modelling with the cost of increasing modelling complexity.

Proposal 5: Consider the entire video stream (I-frames, P-frames etc) to be transported on a bearer with a single associated QoS class.

Proposal 6: RAN1 should set up different models for different kinds of data streams in AR2 UL.

Proposal 7: RAN1 to study layer-1 aspects of large packet transmission with better reliability than eMBB and/or with low packet delay

**LG Electronics**

Proposal 1: Not to include additional XR applications for RAN1 study than XR applications agreed in RAN1#103-e, e.g., VR1/2, AR1/2 and CG.

Observation 1: AR1 and AR2 are essential applications for XR in 5G ecosystem while VR1, VR2 and CG applications can be considered as extension of traditional multi-media services.

Proposal 2: If prioritization for study among XR applications is necessary, AR1/2 should be prioritized over other XR applications.

Proposal 3: Not to have more than one mandatory values per each application for air interface PDB for DL video stream.

* 10 ms for VR/AR and 15 ms for CG

Proposal 4: If a single stream is used for both I-frame and P-frame for DL traffic modelling, a UE is declared a satisfied UE if more than X (%) of packets are successfully transmitted within a given air interface PDB.

* X is [95] which is a single value

Proposal 5: UL Traffic model for video stream/scene update/audio/data

* Frame per second (fps)
  + 60 fps (SA4 input) – no jitter
* Average data rate for UL video stream:
  + 20 Mbps (~half of the average of VR/AR average data rate for DL video stream)
* Truncated Gaussian distribution is used for the packet size distribution
  + Parameters of Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)
    - Mean: Derived from average data rate and fps as follows.
      * (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
    - STD: same as in DL
    - Max packet size: same as in DL
    - Min packet size: FFS (need input from SA4)
* PDB
  + [100] ms (based on SA4 input)

Proposal 6:

* In the case where two streams are used for UL traffic modelling, one for control/pose and the other for video stream/scene update/audio/data, a UE is declared a satisfied UE if more than X1 (%) of packets for control/pose are successfully transmitted within a given air interface PDB1, AND more than X2 (%) of packets for video stream/scene update/audio/data are successfully transmitted within a given air interface PDB2.
  + X1 is [99] and X2 is [95]
  + PDB1 is 10ms and PDB2 is [100]ms

Proposal 7:

* Multi-stream is not further considered in DL for VR1/VR2/CR/AR1/AR2 applications
* Multi-stream is not further considered in UL for VR1/VR2/CG applications
  + FFS whether to support dual-stream for AR application in UL

**InterDigital, Inc.**

Observation 1: For CG there can be more than one traffic flow in UL, consisting of the following traffic types:

* User actions (e.g. gamepad controller, HMD)
* Control data (e.g. protocol flow control, keep-alive messages)

Observation 2: For CG, the UL traffic can be represented by 2 different traffic flows with the following characteristics:

* User Actions
  + Packet arrival is aperiodic and correlated with user activity, where 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, different transport protocols are used for carrying user actions and control data in UL

Observation 4: The sensitivity of QoE to changing QoS is significantly different between the 2 traffic flows

Observation 5: Similar to CG, for VR the UL traffic can be represented by 2 traffic flows

Observation 6: For AR, the UL traffic includes encoded video/media, in addition to user actions and control data

Observation 7: For AR, the traffic characteristics of encoded video/media is significantly different than the other UL traffic flows

Based on these observations, the following conclusions were made:

Proposal 1: RAN1 uses 2 different traffic flows in UL as mandatory for CG evaluations

Proposal 2: The same number of traffic flows in UL used for CG evaluations can also be used for VR evaluations

Proposal 3: RAN1 uses at least 2 different traffic flows in UL as mandatory for AR evaluations. FFS for using more than 2 dfferent traffic flows in UL for AR

It can be further discussed on whether a third traffic flow is considered for AR evaluations.

**AT&T**

Proposal 1: In addition to M1=1 and M2=1 streams, support M1=2 and M2=2, where in both the DL and UL a user has one video stream based on a Truncated Gaussian packet size distribution and one data/control stream based on a fixed packet size and inter-arrival time.

Proposal 2: For both DL and UL consider mixed traffic scenarios with different ratios of UEs with XR and eMBB traffic (e.g. based on FTP Model 3).

**NTT DOCOMO, INC.**

Proposal 1:

* *Consider to study XR conference as optional.*

Proposal 2:

* *Adopt two streams for UL for AR applications*
  + *Traffic model for pose/control information can be same as CG/VR.*
  + *Traffic model for scene update/video/audio data,*
    - *Periodicity: 60 fps*
    - *Data rate: 20 Mbps*
    - *PDB: 60 ms*

Proposal 3:

*RAN1 continues to discuss the statistical models for VR1 and AR1 considering corresponding traffic model discussion in SA4*

# Appendix-B (previous agreements)

## RAN1 #103-e

Agreement: **XR applications**

RAN1 confirms that diverse applications of VR1/2, AR1/2,~~(XR conference FFS),~~ CG are of interest for study. Potential prioritization/down selection of these applications for evaluation is to be discussed after detailed traffic models and relevant evaluation assumptions are stable.

* FFS: other applications, e.g., XR conferencing

Agreement: **Traffic model**

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.

Agreement:

Adopt the following deployment for XR/CG evaluations

* Indoor hotspot: FR1 and FR2
  + Detailed definition of Indoor hotspot refers to TR 38.913.
  + Channel model: InH. Detailed definition of InH refers to TR 38.901.
* Dense urban: FR1 and FR2
  + Detailed deployment refers to TR 38.913, where single layer with Marco layer is assumed.
  + Channel model: UMi. Detailed definition of UMi refers to TR 38.901.

FFS: Whether to prioritize FR1 for evaluation.

Note 1: When selecting the deployment and evaluation assumptions for XR/CG evaluations, it is up to company to evaluate FR1 or FR2 or both for the frequency range.

Note 2: It does not mean that all applications are evaluated for all the deployment scenarios.

Agreement:

Urban Macro can be ~~optionally~~ reported for XR/CG evaluations only for FR1.

* FFS: whether Uma is optional or not
* Following parameters can be assumed.

|  |  |
| --- | --- |
| **Parameter** | **Proposed value** |
| **Urban Macro (FR1)** |
| Layout | 21cells with wraparound ISD = 500 m |
| BS Tx power | FR1: 49 dBm/20 MHz |

Agreement:

It is to be further discussed how to prioritize the combinations of deployment scenarios and applications after traffic models for each application are stable.

Agreement:

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

Agreement:

* Adopt the simulation assumptions in table 1 as below

Table 1: Simulation assumptions for XR evaluation (Part 1) (updated)

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Proposed value** | |
| **Indoor hotspot FR1/FR2** | **Dense urban FR1/FR2** |
| Layout | 120m x 50m ISD: 20m TRP numbers: 12 | 21cells with wraparound ISD: 200m |
| Carrier frequency | FR1: 4 GHz  FR2: 30 GHz | |
| Subcarrier spacing | FR1: 30 kHz  FR2: 120 kHz | |
| BS height | 3m | 25m |
| UE height | hUT=1.5 m | |
| BS noise figure | FR1: 5 dB  FR2: 7 dB | |
| UE noise figure | FR1: 9 dB  FR2: 13 dB | |
| BS receiver | MMSE-IRC | |
| UE receiver | MMSE-IRC | |
| Channel estimation | Realistic  FFS:Ideal(optional) | |
| UE speed | 3 km/h | |
| MCS | Up to 256QAM | |
| BS antenna pattern | Ceiling-mount antenna radiation pattern, 5 dBi | 3-sector antenna radiation pattern, 8 dBi |
| UE antenna pattern | FR1: Omni-directional, 0 dBi,  FR2: UE antenna radiation pattern model 1, 5dBi | |

Agreement:

Adopt the following UE distribution for XR/CG evaluation for outdoor scenario

* For outdoor scenario:
  + FR1: 80% indoor, 20% outdoor
  + FR2: 100% outdoor

Other UE distribution can be evaluated optionally.

Agreement:

Adopt the following TDD configuration for XR/CG evaluation

* FR1:
  + Option 1: DDDSU
  + Option 2: DDDUU
* FR2:
  + Option 1: DDDSU

FFS detailed S slot format

Note: Other TDD configuration or FDD can be optionally evaluated.

Agreement:

Adopt the following BS antenna parameters for indoor scenario for XR/CG evaluation

* FR1:
  + 32 TxRU, (M, N, P, Mg, Ng; Mp, Np) = (4,4,2,1,1;4,4)
  + (dH, dV) = (0.5, 0.5)λ
* FR2:
  + Option 2: 2 TxRU, (M, N, P, Mg, Ng; Mp, Np) = (16, 8, 2,1,1;1,1)
  + (dH, dV) = (0.5, 0.5)λ

Other BS antenna parameters can be optionally evaluated

Agreement:

For XR/CG evaluation, adopt the following assumptions for downtilt

* Dense Urban
  + FFS: 6 or 12 degree
  + ~~Other downtilt can be optionally evaluated.~~
* Indoor hotspot
  + 90° (pointing to the ground)

Other downtilt can be optionally evaluated

Agreement:

* Adopt the simulation assumptions in table 3 as below

Table 3: Simulation assumptions for XR evaluation (Part 3)

|  |  |
| --- | --- |
| **Power control parameter** | Companies should report |
| **Transmission scheme** | Companies should report~~, such as Type I/II codebook, rank assumption~~ |
| **Scheduler** | SU/MU-MIMO PF scheduler (company to report SU or MU),  other scheduler (e.g., delay aware scheduler) is up to companies report |
| **CSI acquisition** | Realistic  Both CSI feedback and SRS are considered  Companies should report  CSI feedback delay, CSI report periodicity, whether using CSI quantization, CSI error model or not,  Assumptions on SRS: periodicity, processing gain, processing delay, etc  and etc. |
| **PHY processing delay** | Baseline: UE PDSCH processing Capability #1  Optional: UE PDSCH processing Capability #2  Companies should report gNB processing delay, e.g. DL NACK to retransmission delay, UL previous transmission to current transmission delay and etc. |
| **PDCCH overhead** | Companies should report |
| **DMRS overhead** | Companies should report |
| **Target BLER** | Companies should report |
| **Max HARQ transmission** | Companies should report |

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.

Agreement:

System bandwidth for XR/CG evaluations are as follows.

* For FR1,
  + Baseline: 100 MHz
  + Optional: 20/40 MHz (FFS: 200 MHz)
* FFS FR2

Agreement:

For outdoor scenarios, the ~~baseline~~ BS antenna parameters are as follows.

* FFS FR1,
  + Option 1: 64 TxRU, (M, N, P, Mg, Ng; Mp, Np) = (8,8,2,1,1;4,8)
  + Option 2: 32 TxRU, (M, N, P, Mg, Ng; Mp, Np) = (8,2,2,1,1,8,2)
  + Option 3: 32TxRUs (M, N, P, Mg, Ng; Mp, Np) = (4,4,2,1,1,4,4)

(dH, dV) = (0.5λ, 0.~~8~~5λ)

* FR2:
  + TxRU, (M, N, P, Mg, Ng; Mp, Np) = (4,8,2,2,2;1,1)

(dH, dV) = (0.5λ, 0.5λ)

Other configurations can be optionally evaluated.

Agreement:

UE antenna parameters for XR/CG evaluations are as follows

* FR1:
  + Baseline: 2T/4R, (M, N, P, Mg, Ng; Mp, Np) = (1,2,2,1,1;1,2), (dH, dV) = (0.5, N/A)λ
  + Optional: 4T/4R, 1T/2R, 2T2R
* FFS FR2: down-selection between the next two options. Please indicate if you have preference.
  + Option 1 (Follow Rel-17 evaluation methodology for FeMIMO in [R1-2007151](file:///E:\Workspace\3GPP%20related\3GPP%20meeting\2021\2021.Q2\RAN1%23104b-e\Summary\Docs\R1-2007151.zip))
    - (M, N, P)=(1, 4, 2), 3 panels (left, right, top)
    - (Mp, Np) is up to company. Need to be reported with simulation result.
  + Option 2 (from TR 38.802 – developed in Rel-14)
    - 4Tx/4Rx: (M, N, P, Mg, Ng; Mp, Np) = (2,4,2,1,2;1,2), (dH,dV) = (0.5, 0.5)λ, the polarization angles are 0° and 90°

Agreement:

BS Tx power for XR/CG evaluations are as follows

* For Indoor hotspot:
  + FR1:
    - 24 dBm per 20 MHz
  + FR2:
    - 23 dBm per 80 MHz. EIRP should not exceed 58 dBm
* For Dense urban:
  + FR1:
    - 44 dBm per 20 MHz
  + FR2:
    - 40 dBm per 80 MHz. EIRP should not exceed 73 dBm

For system BW larger than above, Tx power scales up accordingly.

Agreement:

UE max Tx power for XR/CG evaluations are as follows

* FR1: 23 dBm
* FR2: 23 dBm, maximum EIRP 43 dBm

Agreement: **Baseline power evaluation methodology**

~~If UE power consumption is agreed as a KPI for evaluation of XR performance over NR,~~TR38.840 is the baseline methodology potentially with some modifications if necessary.  RAN1 aim to minimize modeling effort. ~~For example, the following aspects can be considered for further discussion but not limited to.~~

* ~~FFS whether/how to model UE power consumption for UE tx power other than 0dBm and 23dBm,~~
* ~~FFS whether/how to model UE power consumption for UL slots that are not defined in TR38.840~~
* ~~FFS whether/how to model UE power consumption for ‘S’ slot~~
* ~~FFS whether/how to model UE power consumption for 400MHz in FR2 including scaling rule for FR2 BWP adaption.~~
* ~~FFS whether/how to model UE consumption for the corresponding number of Tx antennas~~
* ~~FFS whether/how to model the UE power consumption for UE tx power under FR2~~

Agreement:

* RAN1 continues to discuss evaluation methodologies for UE power consumption and system capacity.
* RAN1 is to discuss whether/how to study/evaluate mobility and coverage at a later stage, e.g., starting from Q1 2021.

## RAN1 #104-e

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
  + 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
* 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

Agreements

* Statistical traffic model for a single DL video stream for a single UE
  + The statistical traffic model for a single UE 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.



* Frame per second (fps) for DL video stream for a single UE
  + 60 fps (baseline)
  + 120 fps (optional)
  + Other values, e.g., 30, 90 fps can be also optionally evaluated.
* Average data rate for DL video stream:
  + VR/AR: 30, 45 Mbps @60fps (baseline)
    - ~~30,~~ 60 Mbps @60fps (optional)
    - Note: this is the aggregated data rate when applicable
  + CG: 8, 30 Mbps @60fps (baseline)
    - ~~8,~~ 45 Mbps @60fps (optional)
  + Other values (in combination with fps) can be also optionally evaluated.
* Truncated Gaussian distribution is used for the packet size distribution of video stream for AR/VR/CG.
  + Other distribution is not precluded.
* (Working assumption) Parameters of Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)
  + Mean: Derived from average data rate and fps as follows.
    - (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
  + STD
    - TBD
  + Max packet size
    - TBD
  + Min packet size
    - TBD
    - FFS whether or not to use this parameter
* 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, e.g., 99, 95
      * FFS different values for I-frame and P-frame if evaluation of them is agreed.
      * Other values can be optionally evaluated
* DL traffic model: video stream
* (Working assumption) Parameters of Truncated Gaussian distribution for Packet size (note: these parameter values are those before the truncation)
  + Mean: Derived from average data rate and fps as follows.
    - (average data rate) / (fps for video stream, i.e., # packets per second in our statistical model) / 8 [bytes]
  + STD
    - [15% of Mean packet size derived above]
    - Note: The above value is an example for further investigation, and is to be revisited potentially with more inputs from companies in RAN1#104-bis-e
  + Max packet size
    - [1.5 x Mean packet size derived above]
    - Note: The above value is an example for further investigation, and is to be revisited potentially with more inputs from companies in RAN1#104-bis-e
  + Min packet size
    - TBD
    - FFS whether or not to use this parameter
    - Note: This is to be revisited potentially with more inputs from companies in RAN1#104-bis-e.
* Jitter for DL video stream for a single UE
  + (Already agreed) Per the agreed statistical traffic model, arrival time of packet k is k/X1000 [ms] + J [ms], where X is the given fps value and J is a random variable.



* + (Newly proposed agreement) J is drawn from a truncated Gaussian distribution:
    - Mean: [0]
    - STD: [2 ms]
    - Range: [[-4, 4]ms]
      * Note: The values ensure that packet arrivals are in order (i.e., arrival time of a next packet is always larger than that of the previous packet)
    - Note: The above values for mean, STD and Range are working assumption for initial simulations, and is to be revisited potentially with more inputs from companies in RAN1#104-bis-e
* Air interface PDB for DL video stream
  + VR/AR:
    - 10ms
    - Other values, e.g., 5ms, 20 ms can be optionally evaluated.
  + CG:
    - 15ms
    - Other values, e.g., 10ms, 30ms can be optionally evaluated.
  + FFS whether or not to have more than one mandatory value

Working assumption: 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 (SA4 input)
  + PDB: 10 ms
* AR
  + FFS

Agreements: On evaluation of multiple streams/flows:

* FFS the following in RAN1#104-bis-e
  + Whether/how to model and evaluate I-frame and P-frame for both DL and UL, e.g., separate definition of fps, packet size, QoS requirements (e.g., PER, PDB), etc.
  + Whether/how to separately model and evaluate two streams of video and audio/data for both DL and UL
  + Whether/how to model and evaluate FOV (high-resolution) and non-FOV (lower-resolution omnidirectional) streams, e.g., separate definition of fps, packet size, QoS requirements (e.g., PER, PDB), etc

Agreement: Adopt following update for TDD configuration for XR/CG evaluation

* FR1:
  + Option 1: DDDSU
  + Option 2: DDDUU
* FR2:
  + Option 1: DDDSU
  + Option 2: DDDUU

Detailed S slot format is 10D:2F:2U. Other S slot format(s) can also be optionally evaluated.

Further clarify that for option 2 for FR1/FR2, there is [2]-symbol gap at the end of third “D” slot of  DDDUU.

FFS whether or not to differentiate the two options (e.g., mandatory vs. optional)

Agreement**:** For XR evaluation, ideal channel estimation can be optionally evaluated.

Agreements**:** System bandwidth for XR/CG evaluations are as follows.

* For FR1,
  + Baseline: 100 MHz
  + Optional: 20/40 MHz, 2\*100 MHz with CA
* FR2
  + Option 1: 100 MHz
  + Option 2: 400 MHz

Companies should report the CA setting if CA is adopted.

Other system bandwidth can also be optionally evaluated.

Agreements**:**For outdoor scenarios, the BS antenna parameters are as

* Option 1: 64 TxRU, (M, N, P, Mg, Ng; Mp, Np) = (8,8,2,1,1;4,8)
* Option 2: 32 TxRU, (M, N, P, Mg, Ng; Mp, Np) = (8,2,2,1,1,8,2)

Company to report the BS antenna parameters for XR/CG evaluation.

Other BS antenna parameters can also be optionally evaluated.

Agreements**:**For FR2, UE antenna parameters for XR/CG evaluations are as follows.

* Option 1 (Follow Rel-17 evaluation methodology for FeMIMO in R1-2007151)
  + (M, N, P)=(1, 4, 2), 3 panels (left, right, top)
* Option 2 (from TR 38.802 – developed in Rel-14)
  + 4Tx/4Rx: (M, N, P, Mg, Ng; Mp, Np) = (2,4,2,1,2;1,2), (dH,dV) = (0.5, 0.5)λ, the polarization angles are 0° and 90°

Company to report the UE antenna parameters for XR/CG evaluation.

Other UE antenna parameters can also be optionally evaluated.

Agreements**:** For XR/CG evaluation, adopt following assumptions for BS height for Urban Macro

|  |  |
| --- | --- |
| **Parameter** | **Proposed value** |
| **Urban Macro (FR1)** |
| BS height | 25m |

Agreements**:** For Dense urban and Urban Macro, the UE height for indoor UEs is updated as following based on Table 6-1 in TR 36.873.

|  |  |  |
| --- | --- | --- |
|  |  | Urban Micro/Macro cell  with high UE density  (3D-UMi) /(3D-UMa) |
| UE height (*hUT*) in meters | general equation | *hUT*=3(*nfl* – 1) + 1.5 |
| *nfl* for outdoor UEs | 1 |
| *nfl* for indoor UEs | *nfl* ~ uniform(1,*Nfl*) where  *Nfl* ~ uniform(4,8) |

Agreements: At least for XR/CG capacity evaluation, for DL and UL

* Baseline: DL and UL performances are evaluated independently
* Optional: DL and UL performance are evaluated together
* FFS details both the baseline and the optional evaluations

Agreements: For Dense urban for XR/CG evaluation, update the agreement in RAN1 #103e for channel model as follows.

* Dense urban: FR1 and FR2
  + Channel model: ~~UMi~~ UMa. Detailed definition of ~~UMi~~ UMa refers to TR 38.901.

Agreements: For XR/CG evaluation, adopt 12 degree for downtilt for Dense Urban in FR1.

* Other downtilt value can also be optionally evaluated

Agreements: To facilitate further discussion on evaluation of power saving effect of different power saving schemes, the following references are defined.

* Case 1 (baseline): UE power consumption assuming UE is always ON, i.e., UE is always available for gNB scheduling.
* Case 2 (FFS optional or baseline): UE power consumption assuming Rel-15/16 CDRX configuration
  + FFS CDRX configuration details
* Company can also optionally evaluate ~~for~~ other cases, e.g.
  + Genie: UE power consumption assuming that UE is in a sleep state (e.g., micro/light/deep sleep as defined in TR38.840) whenever there is neither DL data reception nor UL transmission. From the gNB scheduling perspective, UE is always available for scheduling, i.e., there is no difference from Baseline in gNB scheduling and corresponding UE Tx/Rx. ~~It is noted that Genie is not a power saving scheme but the result may serve as an upper bound of power saving gain of power saving techniques, which may potentially motivate development of new power saving techniques that can approach the Genie performance.~~
  + R15/16/17 power saving techniques for connected mode, e.g., BWP, PDCCH skipping, search space switching, etc.

**Decision:** As per email posted on Feb 5th,

Agreements:

UE power consumption (i.e., power saving gain of the evaluated scheme) for XR is evaluated in conjunction with impact on latency, user experience, and capacity.  In this regard, the following table is used to collect results for system level simulation from companies as a starting point.

* FFS all UEs or only satisfied UEs are included for obtaining the PS gain

Table 1 Evaluation of UE power saving schemes for e.g., {dense urban, AR, FR1}

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Power Saving Scheme | Power Saving Gain (PSG) compared to Case 1 | | | | #satisfied UEs per cell2 / #UEs per cell3 |
| Baseline | Optional | | |
| Mean PS gain | PS gain of 5%-tile UE in PSG CDF1 | PS gain of 50%-tile UE in PSG CDF1 | PS gain of 95%-tile UE in PSG CDF1 |
| Case 1 | - | - | - | - | K1 / N |
| Case 2 | X1 % | Y1 % | Z1 % | U1% | K2/ N |
| Case X | X2 % | Y2 % | Z2 % | U2% | K3 / N |
|  |  |  |  |  |  |

Note 1: CDF of power saving gains of ~~each~~ UE

Note 2: # of satisfied UEs per cell among # of UEs per cell (=N).

Note 3: # of dropped UEs per cell (=N) that needs to be the same for all power saving schemes to be evaluated.

Note 4: company to provide the detailed simulation assumptions including parameter values for each case, e.g. CDRX parameters

~~Note 5: company can report one or more power saving gain metrics (i.e. mean PS gain or PS gain of 5%/50%/95%/-tile UE in PSG CDF) for each power saving scheme~~

Agreements: For UL UE power consumption evaluation for UE with transmit power X [0,23] dBm, adopt the following

* Option 1 ~~(Baseline)~~: Consider only two Tx power values as defined in TR 38.840
  + Power number is given as **A** for X= [0, M)dBm and **B** for X =[M, 23]dBm, where **A** and **B** (defined in 38.840) correspond to power consumption numbers for a given uplink slot for 0dBm and 23dBm respectively.
    - M = [20]
    - Other value(s) of M can be optionally evaluated
  + ~~Companies to provide detailed assumptions on UE power consumption for Tx power values other than 0 and 23 dBm~~ 
    - ~~E.g. Power number is given as~~ **~~A~~** ~~for X= [0, 20)dBm and~~ **~~B~~** ~~for X =[20, 23]dBm, where~~ **~~A~~** ~~and~~ **~~B~~** ~~(defined in 38.840) correspond to power consumption numbers for a given uplink slot for 0dBm and 23dBm respectively.~~
* Option 2 ~~(FFS mandatory or optional)~~: Linear interpolation method in linear scale for Tx power values other than 0 dBm and 23 dBm
* FFS whether or not to differentiate the two options (e.g., mandatory vs. optional)
* FFS whether or not to consider UE with transmit power less than 0 dBm