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

e-Meeting, August 16th – 27th, 2021

Source: Moderator (Qualcomm)

Title: Summary on XR Traffic Model

Agenda Item: 8.14.1

Document for: Discussion and Decision

# Introduction

This contribution is a summary on the email discussion on XR traffic model.

# Outcome of RAN1 #106-e

# Round 2: Traffic model for multi-stream: I-frame and P-frame

## Alpha value

1. **Please share your preference between the next two alternatives for alpha value. If you have a view/proposal other than the two, please feel free to discuss/propose.**
* Alternative 1:
	+ Alpha value: 2.0.  Other values, e.g., 1.5 or 3.0 can be optionally evaluated
		- This alpha value assumption applies to both Option 1A (slice-based) and Option 1B (GOP-based) evaluations
* Alternative 2:
	+ For Option 1A (Slice-based), Alpha value: 2.0.  Other values, e.g., 1.5 or 3.0 can be optionally evaluated
	+ For Option 1B (GOP-based), Alpha value: 1.5.  Other values, e.g., 2.0 or 3.0 can be optionally evaluated

|  |  |
| --- | --- |
| **Company** | **Comment** |
| vivo | We prefer Alternative 2. Based on the P-traces profile provided by SA4, the larger alpha value (e.g., 2.0, 3.0 or even higher) should be applied for Option 1A (Slice-based). And for Option 1B (GOP-based), the alpha values are relatively low (e.g., 1.5). It also can be obtained from theoretical analysis that P-slice is the leading factor affecting capacity for slice-based case, whereas the I frame is the leading factor affecting capacity for GOP-based. Therefore, due to the different traffic characteristics of slice-based and GOP-based, using a unified alpha value for such two models (i.e., Alternative 1) is not necessary. |
| LG | We slightly prefer Alt 2 if it represents the each of the encoding schemes better.Note that assuming different baseline alpha values per each Option as in Alt 2 does not increase the number of simulation runs. |
| MTK | We prefer Alt. 1 to have a more unified assumption but we are also fine with Alt. 2 if that is the majority view. |
| Ericsson | We prefer Alt.1. It reduces the number of evaluation cases.  |
| Huawei, HiSilicon | We prefer Alt 1.We do appreciate companies’ deep analysis and constructive comments on this point. However, since there could be many codec configurations in real applications, the value of alpha could be quite diverse. For example, as analysed by quite a few companies, based on the SA4 traces, the value range of alpha is 1~2. Meanwhile, in SA4 LS to SA2/RAN1 in R1-2102308, it is mentioned that*“… intra-coded information typically requires 3 to 6 times the amount of data compared to inter-coded information…”*. So it seems alpha value can also take values from 3 to 6, and SA4 does not mention “3 to 6 times” is for Option 1A (Slice-based) or for Option 1B (GOP-based) in this LS.So in general, we agree different codec configurations may have different alpha values. However, whether there are any typical configurations or which configurations are more typical is beyond RAN1’s expertise. So we suggest RAN1 does not need to do so detailed differentiations. We support Alt 1 to have a unified assumption and simplify further evaluations/discussions, and companies can anyway choose to simulate other values based on their interest. |
| ZTE, Sanechips | Alt 1. Prefer a single value as analysed in R1-2106526 |
| Xiaomi | We prefer to Alt 1 to simplify the evaluation. |
| QC | We prefer Alt 1. We think Alt 1 could be a good starting point which captures well the fact that I frame is larger than P frame. Moreover, Alt 1 supports other values optionally, so, in case companies want to evaluate different value for Sliced-based/GOP-based model, they can do that.Given that the choice of alpha depends on video encoding/transfer algorithm (which is beyond the RAN1 scope), it seems that it is reasonable/safe to start with a single baseline value w/ optional values (Alt 1) rather than assuming different values for GOP-based and slice-based with alpha\_GOP < alpha\_sliced (Alt 2).It is not well understood in RAN1, whether/why the alpha values are different for GOP-based and sliced-based method, and if different, whether alpha\_GOP < alpha\_sliced or the other way around. |
| OPPO | Alt-2. Agree with reasoning in LG’s comment. Does Alt-1 mean there is effectively no difference at all from evaluation perspective between sliced-based and GOP-based?  |
| CATT | We prefer Alt 1. Alt 1 is a good staring point for generalized traffic model. The XR contents would change the alpha value.  |
| DOCOMO | We prefer Alt.1. With Alt.1, companies can also provide evaluation results with alpha value = 1.5 for GOP if they prefer to do it. |
| InterDigital | We are ok with Alt 1 as a starting point and to simplify evaluation assumptions.  |
| vivo2 | In accordance with the companies’ analyses above, we would like to provide some further comments below.First of all, we don’t agree with the comment that a unified alpha will simplify evaluation assumptions. On the contrary, we think the unified alpha does not reflect the difference by adopting different codec configurations in real applications. In spite of typical configurations for GOP-based and slice-based is beyond RAN1 knowledge, but we still be able to identify the alpha difference between GOP-based and slice-based from SA4’s inputs. Secondly, as analysed in our contribution [R1-2106629], based on the SA4 LS in [R1-2102308], we can find that “intra-coded information typically requires 3 to 6 times the amount of data compared to inter-coded information”. Besides, for the GOP-based traffic model, by analyzing the S-trace data of VR2-5 provided by SA4, the average size ratio between I-frame and P-frame is around 1.15, while the range of the ratio is from 0.86 to 7.95. **Hence, the value of 1.5 is a reasonable value, at least for GOP-based. This is why we prefer to define alpha=1.5 as the GOP-based baseline.** Figure 3. The CDF of ratio between the sizes of I-frame and P-frame based on the S-trace data. |
| Apple | It seems proponent companies for both alternatives raise some good points, e.g. vivo’s input according to SA4’s LS. |

##  [PER\_I, PER\_P, PDB\_I, PDB\_Q] values

1. **Please share your preference between the next two options for values of [PER\_I, PER\_P, PDB\_I, PDB\_Q]. If you have a view/proposal other than the two, please feel free to discuss/propose.**
* Option 1: RAN1 agree only upon the above reference case, while leaving other study cases up to companies.
	+ Reference case
		- For DL
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1 %, 10ms, 10ms] for AR/VR
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1 %, 15ms, 15ms] for CG
		- For UL AR aggregating streams of scene, video, data, and audio
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1 %, 30ms, 30ms]
* Option 2:
	+ Set of cases to be evaluated is completely left to companies (i.e., RAN1 will not discuss further common set of evaluation cases)

|  |  |
| --- | --- |
| **Company** | **Comment** |
| vivo | We are OK with both Option 1 and Option 2, and Option1 is slightly preferred, since the PDB and PER values of Option1 are in consistent with single stream traffic model and current 5GS operation. Furthermore, reference cases can also adopt the optional evaluation values from single stream traffic model.  |
| LG | As what we are discussing is basically optional, Option 2 should be fine and is our preference. But, if a majority wants to have some reference case to facilitate comparison, we can live with Option 1. |
| MTK | Option 1. Reference cases are required as a baseline, and companies can provide possible enhancement directions (ex. manipulation with PER/PDB) compared with the baseline. |
| Ericsson | Option 1. We already now have too many evaluation assumptions. |
| Huawei, HiSilicon | We support Option 1 with some modifications.If such PER/PDB values are totally left to company report, different companies may choose quite different values. As a consequence, the results from different companies are not comparable, and it’s not easy for RAN1 do draw conclusions. Therefore, agreeing some sets of PER/PDB can help RAN1 discussion and is preferred. Quite a few companies have already identified the following two essential aspects of I/P-frame:* Due to the difference of importance, I-frame could have lower PER than P-frame
* Due to the difference of decoding delay, I-frame could have larger PDB than P-frame

So we think it’s important to evaluate the cases where I/P-frame could have different PER/PDB values, which can help RAN1 gain more insight on such aspects. For example, [PER\_I, PER\_P, PDB\_I, PDB\_P] = [0.5 %, 5 %, 17ms, 9ms] for AR/VR could be considered, which was proposed by FL in the 1st round.In 1st round discussions, we also suggested the following PER/PDB values (take DL AR/VR as an example). The motivations is control the number of variables when we do simulations, so that companies can better know the impact of changing one specific parameter on capacity. * Case 2: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [0.5%, 5%, 10ms, 10ms]
* Case 3: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1%, 17ms, 9ms]
* Case 4: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 5%, 10ms, 10ms]
* Case 5: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1%, 17ms, 10ms]
* Case 6: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 5%, 17ms, 10ms]

In general, the intention of proposing the above cases is to help RAN1 discussion on gaining more insight on I/P-frame, and also to avoid the situations of very diverse simulation cases from companies. We are open for discussion on the detailed values.If companies have strong concerns on agreeing some additional sets of PER/PDB values in this meeting, we might be ok to leave them to company report. But we suggest to include at least one case as an example as shown below, which is mainly to remind companies to consider the essential aspects of I/P-frame, i.e., different PER/PDB requirements. We assume this will not be very controversial since it is anyway just an example.Btw: a small typo in title of Section 3.2, i.e., “PDB\_~~Q~~P”.In summary, we suggest the following changes in red.== * Option 1: RAN1 agree only upon the ~~above~~below reference case, while leaving other study cases up to companies.
	+ Reference case
		- For DL
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1 %, 10ms, 10ms] for AR/VR
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1 %, 15ms, 15ms] for CG
		- For UL AR aggregating streams of scene, video, data, and audio
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1 %, 30ms, 30ms]
	+ Other cases, i.e., other values of [PER\_I, PER\_P, PDB\_I, PDB\_P] can be evaluated and up to company report, e.g.
		- For DL
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [0.5 %, 5 %, 17ms, 9ms] for AR/VR
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [0.5 %, 5 %, 20ms, 14ms] for CG
		- For UL AR aggregating streams of scene, video, data, and audio
			* [PER\_I, PER\_P, PDB\_I, PDB\_P] = [0.5 %, 5 %, 40ms, 27ms]
 |
| ZTE, Sanechips | Ok for option 1, though prefer differentiation between I/P component. |
| Xiaomi | We are fine with both option 1 and option 2, and slightly prefer to option1 as the evaluation results from companies can have more aligned assumption. However, as there is still different opinions on the purpose to do these evaluations in RAN1, we do not think further proposal on study cases is needed. |
| QC | We support Option 1. RAN1 can start with a single reference configuration and companies can consider different pairs of values. Given that the performance requirements of (PER, PDB) pair needs to be determined based on upper layer XR user experience, it would be a bit arbitrary for RAN1 to determine a few fixed values for evaluation. |
| OPPO | We are ok to both options.  |
| CATT | We are OK with either Option 1 and 2. Our preference is Option1 to have single target of reference case.  |
| DOCOMO | We support Option 1 to have single configuration. |
| InterDigital | We are ok with the parameters in Option 1 as a reference case. |
| Apple | Both Options are fine to us.  |

# Round 1: Discussion on open issues

The following is a RAN1 agreement w.r.t. evaluation of two streams for DL.

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

## Traffic model for multi-stream: I-frame and P-frame

In RAN1#105-e, the detailed traffic model for two streams of I-frame and P-frame for DL video was agreed as captured below. The average size ratio between one I-frame/slice and one P-frame/slice and the PER/PDB for each of the I-stream and P-stream are remained as FFS.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Agreement:For the optional evaluation scenario, two streams of I-frame and P-frame for DL video stream (option 1), the traffic models described in the below table are assumed. * FFS: Parameter values of , A, B, C, D, E, F, G, H
	+ Including the possibility of using multiple set of parameter values
* For companies who are evaluating this option, it is recommended to evaluate at least the following scenario: AR/VR, 30Mbps, Dense Urban for FR1 and InH for FR2. It is encouraged to evaluate additional baseline/optional scenarios/configurations.

|  |  |  |
| --- | --- | --- |
| **Two data streams, i.e. M1 = 2** | **Option 1A: slice-based** | **Option 1B: GOP-based** |
| I-stream | P-stream | I-stream | P-stream |
| **Packet modelling** | Slice-level | Frame-level |
| **Traffic pattern** | Both streams are periodic at 60 fps with the same jitter model as for single stream.  | Follow the GOP structure, where GOP size K = 8 with the same jitter model as for single stream. |
| **Number of packets per stream at a time** | 1 | N-1 | 1. frame: 1 or 0

P-frame: 0 or 1At each time instant, there is either only one I-stream packet or only one P-stream packet |
| N = 8: the number of slices per frame. |
| **Average data rate per stream** | e1 | e2 |   |   |
| * R: average data rate of a single stream video
* : average size ratio between one I-frame/slice and one P-frame/slice, e.g. = 1.5, 2, 3
 |
| **Packet size distribution** | Truncated Gaussian distribution |
| Mean =  | Mean =  | Mean =  | Mean =  |
| * [STD, Max, Min]: [10.5, 150, 50]% of Mean packet size
* FPS is the frame rate of the single stream video
 |
| **PER, PDB** | [PER\_I, PER\_P] = [A %, B %][PDB\_I, PDB\_P] = [C ms, D ms] | [PER\_I, PER\_P] = [E %, F %][PDB\_I, PDB\_P] = [G ms, H ms] |

 |

Companies’ views presented in contributions for RAN1#106-e are summarized below.

*Alpha value:*

* 1.5: China Telecom, vivo (for GOP-based), MTK
* 2: HW, ZTE, Samsung, China Telecom
* 3: CATT, Ericsson, vivo (for slice-based), MTK
* Intel: 1.07-1.90

*Other parameters*

* HW: 5 cases for DL and 5 cases for UL
* ZTE: same PDB, PER\_I (1%) > PER \_P (10%)
* Vivo:
	+ Sliced-based: [PER\_I, PER\_P] = [1 %, 5 %], [PDB\_I, PDB\_P] = [5 ms, 10 ms]
	+ GOP-based: [PER\_I, PER\_P] = [1 %, 5 %], [PDB\_I, PDB\_P] = [10 ms, 15 ms, 20 ms]
* Samsung: PDB (10ms for AR/VR, 15ms for CG) and PER (1%)
* China Telecom
	+ [PER\_I, PER\_P] = [1 %, 10 %] or [0.5%, 5%] as baseline
	+ [PDB\_I, PDB\_P] = [10ms, 10ms] or [15ms, 15ms] or [20ms, 20ms ] as baseline
* MTK
	+ [PER\_I, PER\_P] = [1 %, 1 %] or [0.5%, 5%] as baseline
	+ [PDB\_I, PDB\_P] = [10ms, 10ms] and [17ms, 9ms] as baseline
* Ericsson
	+ PDB (10ms for AR/VR, 15ms for CG) and PER (1%)
1. **FL proposals based on RAN1#106-e contributions are given below. Please share your view on these proposals.**
* Alpha value: 2
* [PER\_I, PER\_P] and [PDB\_I, PDB\_P]
	+ For DL
		- Reference case:
			* [PER\_I, PER\_P] = [1 %, 1 %]
			* [PDB\_I, PDB\_P] = [10ms, 10ms] for AR/VR and [15ms, 15ms] for CG
		- Common study case
			* [PER\_I, PER\_P] = [0.5 %, 5%]
			* [PDB\_I, PDB\_P] = [17ms, 9ms] for AR/VR and [20ms, 14ms] for CG
	+ For UL AR aggregating streams of scene, video, data, and audio
		- Reference case:
			* [PER\_I, PER\_P] = [1 %, 1 %]
			* [PDB\_I, PDB\_P] = [30ms, 30ms]
		- Common study case
			* [PER\_I, PER\_P] = [0.5 %, 5 %]
			* [PDB\_I, PDB\_P] = [40ms, 27ms]
* For multi-stream evaluations, a UE is declared as a satisfied UE if each stream meets the PER and PDB requirements, i.e., more than a certain percentage of packets are successfully transmitted within a given air interface PDB.
* Note:
	+ What to study and potentially what to be captured in the TR: Study the impact on capacity from different PDB and PER values for I-frame and P-frame by comparing capacity results between the reference case and study case(s).
	+ The main intent to define a single common study case is to have more results from companies with the same parameter values.
	+ Companies can submit results for more study cases with different parameter values.
	+ We don’t intend to directly compare capacity results (i.e., capacity numbers) for two-stream cases and those for single-stream cases.

|  |  |
| --- | --- |
| **Company** | **Comment** |
| OPPO | We would like to get a clarification of reason behind PER setting: [1%, 1%] for reference and [0.5%, 5%] for study case. We think a reasonable setting is to make the reference case and study case have the user experiences on a full video frame as close to each other as possible. But if we apply a simple comparison, the situation seems in different way. The detail is given below. Assume eI and eP represent the PER of I-frame and PER of P-frame, respectively. Then the total video traffic that is successfully going through for user experience is:  pkt\_sizeI \* (1-eI) + (N-1)\*pkt\_sizeP \* (1-eI) \* (1-eP) = pkt\_sizeP \* (1-eI) \*[+(N-1)\*(1-eP)]We assume this successful video packet volume can somehow link to user experience upon the full video quality. With and N=8, and omitting the common multiplication term, * For [1%, 1%], we have 0.99\*[2+7\*0.99] = 8.841
* For [0.5%, 5%], we have 0.995\*[2+7\*0.95] = 8.607

Therefore, the setting of [0.5%, 5%] leaves UE with less amount of successful video data. In order to have the same amount of successful video data, * the PER of P-frame cannot be larger than 2.29% (when PER of I-frame is down to 0).
* Under PER of I-frame equal to 0.5%, the PER of P-frame needs to be no larger than 1.65%

The above analysis may be too mathematical, but we would like to know the reason to assume 5% PER for P-frame.  |
| Vivo | Although we understand defining common study case is helpful for aligning the results from different companies, why to adopt these PER or PDB values for the common study case should be clarified for reaching consensus. In our views, the values of PER and PDB for multi-stream are supposed to be not only reasonable and feasible, but also be able to provide insight from the evaluation. Given this fact, we suggest to contain multiple values for the common study case.Firstly, the value of alpha affects the capacity performance. In order to investigate the impact of alpha, other alpha value e.g., alpha =1.5 or 3 should also be considered in addition to alpha= 2. Secondly, at least the same PER\_I for common study case with that of reference case needs to be considered. Thirdly, we think at least the same PDB\_P for common study case with that of reference case needs to be considered. It is not clear to us why the value of PDB\_P for common study case is reduced compared to the PDB\_P value in reference case since the impact on capacity performance by reducing PDB\_P is unaware despite the packets size of P frame is smaller than I frame. On contrary, relaxing PDB of P-frame can help to improve the capacity performance. |
| LG | For the suggested alpha value = 2, we think this seems to be rather an extreme case that can be observed only in the case of Option 1A (slot-based), which is fine for us to check the impact on the capacity in this case. For the PERs and PDBs, we don’t a strong view. But, we prefer to focus on evaluations given the KPIs set according to the guidance from higher layer rather than trying to provide ways to increase the capacity by changing the PERs and PDBs. In view of that, we could agree on the reference case, and leave the other cases up to companies to report if it is not easily converged.We would like to understand the numbers in the FL’s proposal. For the PDBs for the common study case, the following values were proposed. • [PDB\_I, PDB\_P] = [17ms, 9ms] for AR/VR and [20ms, 14ms] for CGAs we understand it, [17, 9] tries to match the average PDB per frame in the case of Option 1B (GOP-based). If we want to keep this principle for other cases, for example, [20, 14] could be [22, 14]. And the principle doesn’t seem to hold for Option 1A (slice-based). So, before agreeing on the specific values, some clarifications would be helpful for us. |
| Ericsson | Support.  |
| Huawei, HiSilicon | **On alpha value**: we support 2 as baseline. Other values, e.g., 1.5 and 3, can be optionally evaluated.**On PER/PDB**: We suggest to use a quadruple fashion, e.g., [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1%, 1%, 10ms, 10ms] to avoid any confusion. Otherwise, it’s not very clear whether companies can combine [PER\_I, PER\_P] in reference case and [PDB\_I, PDB\_P] in common study case to get a new case.We suggest Case 1~6 in the following updated proposal, the motivation and physical meaning is as follows:* Case 1 is the reference case (same as FL’s proposal), where I/P frame has the same PER and PDB
* Due to the difference of importance, I-frame could have lower PER than P-frame
	+ This is reflected in Case 2, where PER of I-frame is lower than reference case, and PER of P-frame is higher than reference case
* Due to the difference of decoding delay, I-frame could have larger PDB than P-frame
	+ This is reflected in Case 3, where PDB of I-frame is larger than reference case, and PDB of P-frame is smaller than reference case
* Additionally, we propose Case 4/5/6, this is mainly to control the number of variables when we do simulations, so that companies can better know the impact of changing one specific parameter on capacity
	+ For example, when comparing Case 1 and Case 4, only PER\_P changes, so companies can know the impact of changing PER\_P on capacity
	+ The logic of Case 5, 6 is similar

In addition, we’d like to make the following clarifications * Note1: for CG, maybe we can simply say PER\_I, PER\_P is same as PER\_I, PER\_P of AR/VR, respectively, and PDB\_I, PDB\_P is 5ms longer than PDB\_I, PDB\_P of AR/VR, respectively. This is to reuse the principle of single stream.
* Note2: 6 cases does not mean 6 times of simulations
	+ Because if PDB remains the same, we can get capacity results corresponding to different PER values in a single simulation. It’s just some simple data processing work.
	+ For example, the results of Case 1, 2 can be obtained from a single simulation.
* Note3: we suggest to agree on these cases, instead of leaving them to company report
	+ Because if we only agree on 1 or 2 cases, and leave other cases up to company report, then different companies may choose quite different values. As a consequence, the results from different companies are not comparable, and it’s not easy for RAN1 do draw conclusions.
	+ So agreeing on the cases below in fact reduces the workload and helps RAN1 progress. We are open to discuss the detailed values if companies have concerns
* Note4: we think there is no need to discuss the corresponding user experience of each case
	+ Because such discussions are beyond RAN1’s expertise. And RAN1 also does not discuss the corresponding user experience for single stream case.
	+ So RAN1 is just going to evaluate and report the corresponding capacity for each case. Interested people, e.g., application layer guys, can have their own understanding on such results.

**On the Note**: on the first sub-bullet, we think it’s also meaningful to study and capture the impact on capacity from scheduler which considers I/P-frame characteristics. For example, since I-frame is more important, it’s straightforward that prioritizing transmitting of I-frame can increase capacity, which has already been observed from our simulation results. The idea is similar to delay-aware scheduler, which is simulated by some companies and capacity gain is also observed. Similarly, in power evaluation, some companies also proposed some new power saving schemes and observed gains. Generally, we think such aspects need to be studied and captured in the TR, which aligns with the ideas of delay-aware scheduler and new power saving schemes.In summary, we support the following red changes on the proposal:==* Alpha value: 2 is baseline, 1.5 and 3 are optional
* [PER\_I, PER\_P, PDB\_I, PDB\_P]
	+ For DL
		- AR/VR
			* Case 1: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1%, 1%, 10ms, 10ms]
				+ Reference case
			* Case 2: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [0.5%, 5%, 10ms, 10ms]
			* Case 3: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1%, 17ms, 9ms]
			* Case 4: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 5%, 10ms, 10ms]
			* Case 5: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1%, 17ms, 10ms]
			* Case 6: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 5%, 17ms, 10ms]
		- CG
			* PER\_I, PER\_P: same as PER\_I, PER\_P of AR/VR, respectively
			* PDB\_I, PDB\_P: 5ms longer than PDB\_I, PDB\_P of AR/VR, respectively
	+ For UL AR aggregating streams of scene, video, data, and audio
		- AR/VR
			* Case 1: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1%, 1%, 30ms, 30ms]
				+ Reference case
			* Case 2: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [0.5%, 5%, 30ms, 30ms]
			* Case 3: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1%, 40ms, 27ms]
			* Case 4: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 5%, 30ms, 30ms]
			* Case 5: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 1%, 40ms, 30ms]
			* Case 6: [PER\_I, PER\_P, PDB\_I, PDB\_P] = [1 %, 5%, 40ms, 30ms]
* For multi-stream evaluations, a UE is declared as a satisfied UE if each stream meets the PER and PDB requirements, i.e., more than a certain percentage of packets are successfully transmitted within a given air interface PDB.
* Note:
	+ What to study and potentially what to be captured in the TR: Study the impact on capacity from different PDB and PER values for I-frame and P-frame by comparing capacity results between ~~the reference case and study~~ different case~~(~~s~~)~~, the impact on capacity from scheduler which considers I/P-frame characteristics, etc.
	+ The main intent to define a single common study case is to have more results from companies with the same parameter values.
	+ Companies can submit results for more study cases with different parameter values.
	+ We don’t intend to directly compare capacity results (i.e., capacity numbers) for two-stream cases and those for single-stream cases.
 |
| MTK | We support the FL proposal.We share the same understanding as LG that the proposed value (Ex. [17, 9]) is to match the average PDB per frame and see the performance impact of delay budget adaptation, hence we also can accept [22, 14] suggested by LG instead of [20, 14]. For Option 1A (slice-based), it may require some further study as mentioned by LG, we are open to hear companies’ view on possible settings. For the value of alpha, we prefer to have one value as baseline (1.5 or 2 or 3 are all acceptable to us), while other values can be optionally evaluated for multi-video-stream scenario, to manage the simulation efforts.For vivo’s comment:* It is not clear to us why the value of PDB\_P for common study case is reduced compared to the PDB\_P value in reference case

the intention is to match the average PDB per frame and see the performance impact of delay budget adaptation. We are also open for another evaluation case which has a larger value of PDB\_P for a more complete study on delay budget adaptation (A direct suggestion on the candidate values from vivo would be helpful).For OPPO’s analytical formula, we think it is quite interesting. We are open to set PER\_I/PER\_P to have the same amount of successfully transmitted video data. A direct suggestion on the candidate values would be helpful for the group to choose some suitable settings, and maybe replace the current [0.5%, 5%] candidate.For Huawei’ suggestion on including more cases for simulation, we are open for this suggestion to have a more completed analysis on performance impact of PDB/PER adaptation. If the simulation results for basic scenarios of XR are well-captured in this meeting, then we may have more simulation capacity released to include more cases here. |
| Xiaomi | We have similar questions on how the values of PDB and PER are determined in the proposal. To compare the evaluation results between reference case and common study case, some basic assumption may be necessary. For example, assuming that the XR quality is the same for both cases, or both cases have the same average PDB (as mentioned by LG), or both cases have the same total successful traffic volume (as mentioned by OPPO), etc. Without such kind of assumptions, we are not sure whether the proposed value sets can provide meaningful results. Considering the difficulty to evaluate XR quality in RAN1, we are fine to assume both cases have the same average PDB, and both cases have the same total successful traffic volume (i.e. same average PER).  |
| InterDigital | We are ok with FL’s proposal, in particular the KPI for multi-stream evaluations and the accompanying notes. |
| AT&T | We are OK with the FL proposal. We agree that multi-stream and single-stream capacity numbers cannot be directly compared, however in practice a system will contain a mix of users with different traffic types and service requirements, so some general observations overall about the relative/aggregate impact of different traffic models on the system performance would be useful to capture in the TR at the end of the study. |
| Nokia, NSB | We support the current proposal. We also support the single value for alpha since that is the part of the model itself, increasing the number of values for alpha just increases the amount of traffic models, which is already extensive.  |
| CATT | We are OK with moderator’s proposal. We don’t need to make it complicated with different cases.  |
| Intel | On the value of alpha, we are OK with moderator’s proposal, alpha value: 2.On PER/PDB, the “Reference case” is not defined – is it simply a case obtained by using the same PER/PDB values as a single stream case (it’s better to define it). The principle of determining unequal PER in a way to approximate single stream PER by appropriately weighting the 2 streams is motivated by Option 1B, justification for Option1A is not clear. On the other hand, the motivation for unequal values for PDB is not clear to us. It’s not clear that a larger packet size can justify a larger PDB.If we don’t have sufficient data to derive such values, one option is to define the reference case and the unequal value case can be left to companies’ discretion. |
| QC | Alpha value**:** we think 2 could be the baseline. However, considering the frame sizes / numbers of frames for I and P frame are encoder parameter dependent, other values could be optionally evaluated.In general, we are fine with the framework of having reference case and common study case for efficient / focused study.Here are a few points we want to make.1. As discussed above, current RAN1 evaluation framework is not enough to capture the impact of lower layer performance metrics to upper layer XR user experience. So, when we are interpreting results or capturing observations, **we should avoid making any conclusion assuming final XR user experience be the same across different cases; this is unknown to RAN1.**
2. We see that some companies argue that P frame takes longer time for decoding than I frame. I and P frame certainly take time for decoding, however, from our understanding so far, they are short enough and **we see don’t see strong reason to differentiate PDB**.
3. The choice of PER needs to be, in theory, determined based on final XR user experience. However, since we don’t have right framework to choose that, we think multiple combination of PERs for I and P could evaluated. When comparing those combinations, we need to be careful. Again, we think the **comparison among different combinations may not be much meaningful (in terms of XR user satisfaction) since we have no idea how those relaxation affect final XR user experience**. We think scheduler-based differentiation of non bottle neck frame type and capacity increase from there could be meaningful to capture.
4. **No comparison between single stream and multi stream** results should be allowed.
 |
| ZTE,Sanechips | In our opinion, firstly, the reference case and the study case in FL summary should be clarified to be suitable in GoP-based traffic model, or Slice-based traffic model, or both of them.Secondly, for the study case, it seems that only the case PDB when I-stream is looser than that of P-stream is simulated. But the cases when I-stream and P-stream have same PDB or I-stream has more stringent PDB than that of P-stream are not included in the common study case. To this regard, we suggest to add some optional common study case with different cases to fully evaluate the impact of different PDB and PER. For example, * Common study case 1
* [PER\_I, PER\_P] = [1 %, 5%]
* [PDB\_I, PDB\_P] = [20ms, 20ms] for AR/VR and [30ms, 30ms] for CG
* Common study case 2
* [PER\_I, PER\_P] = [1 %, 1%]
* [PDB\_I, PDB\_P] = [5ms, 10ms] for AR/VR and [10ms, 20ms] for CG

Thirdly, we wonder if some enhancement schemes are considered for multiple stream transmission, which case is the baseline for enhancement, the reference case or the common study case? |
| DOCOMO | We are OK with the FL proposal. |

## Traffic model for multi-stream: DL video and audio/data

|  |
| --- |
| Agreement:In addition to single stream per UE in DL which is baseline, two streams can be optionally evaluated for DL* Option 1: I-frame + P-frame
	+ Option 1A: slice-based traffic model
	+ Option 1B: Group-Of-Picture (GOP) based traffic model
* Option 2: video + audio/data
* Option 3: FOV + omnidirectional stream
 |

Proposal for Option 2 from Apple and NTT-DOCOMO.

For DL traffic model Option 2, the audio/data flow is efined with:

* A stream aggregating streams of audio and data
	+ Periodicity: 10ms
	+ Data rate: 0.756 Mbps/s or 1.12 Mbps
	+ Packet size: determined by periodicity and data rate
	+ PDB: 30ms
1. **Please share your view on the following proposals from Apple and NTT-DOCOMO.**

For DL traffic model Option 2, the audio/data flow is as follows:

* A stream aggregating streams of audio and data
	+ Periodicity: 10ms
	+ Data rate: 0.756 Mbps/s or 1.12 Mbps
	+ Packet size: determined by periodicity and data rate
	+ PDB: 30ms

|  |  |
| --- | --- |
| **Company** | **Comment** |
| Apple | We support the modelling parameters for DL traffic model Option 2 as given in Question 2. |
| OPPO | We support the Option-2 modeling proposed by Apple and DoCoMo.  |
| DOCOMO | We support the listed parameters above for Option 2. |
| Facebook | We support the listed parameters for Option 2 from Apple and DCM.  |
| Vivo | In our point of view, there is no need to discuss the detailed traffic model parameters for DL audio/data flow, since it should be consistent with the endorsed UL streams of audio and data in RAN1 #104bis-e as below: Agreement:For evaluations of AR in UL:…* Option 3 (Optional): Three streams as defined below
	+ Stream 1: pose/control
		- Traffic model and QoS parameters are same as for pose/control for UL CG/VR.
	+ Stream 2: A stream aggregating streams of scene and video
		- Packet size: Truncated Gaussian distribution with the parameter values same as for DL
		- Periodicity: 60 fps
			* Jitter (optional): same model as for DL
		- Data rate: 10 Mbps (baseline), 20 Mbps (optional)
		- PDB: [60] ms (baseline), [10/15] ms (optional)
	+ Stream 3: A stream aggregating streams of audio and data
		- Periodicity: 10ms
		- Data rate: 0.756 Mbps/s or 1.12 Mbps
		- Packet size: determined by periodicity and data rate
		- PDB: 30 ms

…Besides, it seems the PER requirement for the audio/data flow was not defined, can someone please clarify it? |
| LG | We are fine with the proposals for DL traffic model Option 2 with the understanding that it is reusing the UL traffic model and is based on the SA4 traffic model.In response to vivo’s question. We have agreed on the yellow highlighted part for evaluations of AR in UL:Agreement:For evaluations of AR in UL:* Option 1 (Baseline for power and capacity evaluations): Two streams as defined below
	+ …
* Option 2 (Optional for power evaluation and baseline for capacity evaluation): Single stream as efined below
	+ …
* Option 3 (Optional): Three streams as defined below
	+ …
* Option 4 (Optional): Three streams as defined below
	+ …
* Note: Above PDB values in [ ] for Stream 2 in Option 1 and 3, and Option 2 are to be further discussed and potentially confirmed in RAN1#105-e, where other values can be also discussed if needed.
* In case multiple steams are evaluated for UL AR, a UE is declared as satisfied only when each stream meets the requirement that X (%) of packets are successfully delivered within a given air interface PDB.
	+ X value for pose/control: follow X values for pose/control for CG/VR
	+ X value for other stream: follow X values for DL video stream.

So, unless agreed otherwise, the PER for DL video stream may be reused. |
| Ericsson | Support. OK to reuse the PER as suggested by LG. |
| Huawei, HiSilicon | We are generally ok. |
| MTK | We support the FL proposal. We also think the PER value for audio can be specified explicitly using the suggestion from LG to make this evaluation assumption more self-contained. |
| Xiaomi | We are fine with the proposal. |
| SONY | Support the above proposal. |
| InterDigital | We support the proposal |
| AT&T | We support the proposal for the Option 2 parameters. |
| Nokia, NSB | We support the proposed values. |
| CATT | We are OK with the proposal |
| Intel | We support the proposal. |
| QC | We are fine with FL proposal but PER requirement can be clarified. It could be explicitly given or determined to follow DL video. |
| ZTE, Sanechips | We agree with FL proposal. And we think PER for audio should be clarified. Moreover, according to TR38.824, the audio/data traffic model is similar to the traffic model for transport industry. Therefore, we suggest to add optional PDB and PER for traffic model audio/data to include some audio/data traffic with high reliability and low latency services.In summary, we support the following changes in red on the proposal:For DL traffic model Option 2, the audio/data flow is modeled as follows:* A stream aggregating streams of audio and data
	+ Periodicity: 10ms
	+ Data rate: 0.756 Mbps/s or 1.12 Mbps
	+ Packet size: determined by periodicity and data rate
	+ PDB: 30ms (baseline), 7ms (optional)
	+ PER: 1% (baseline), 0.001% (optional)
 |
|  |  |

## Others

1. **Please share your view on other topics/issues related to XR traffic model if any.**

|  |  |
| --- | --- |
| **Company** | **Comment** |
| OPPO | We propose in our contribution to consider packet size correlation between adjacent video frame (at least for single stream). SA4 did not say whether the sizes of adjacent video frame packet should be correlated or independent, but it is well-known in industry that they are correlated. If RAN1 does nothing in correlation study, the packet size would be independent over time, which has no backup from SA4 and is somehow conflicting to popular understanding in industry.  |
| MTK | We are open to discuss packet size correlation between adjacent video frames suggested by OPPO. However, given that this may require a large re-simulation efforts for all the simulation results currently reported, maybe RAN1 has to make decision carefully. |
| InterDigital | Regarding the evaluation of multiple streams for AR in UL, we propose to consider evaluating another 3 streams option (Option 5), where Stream 3 corresponds to user actions. We think there are benefits in evaluating 3 streams for AR including aggregated video, pose/control and user actions streams, mainly for highlighting the impact on capacity and power consumption when supporting different non-video traffic streams with distinct traffic characteristics (e.g. pose/control stream is periodic, user action stream is aperiodic). The parameters that can be considered for evaluating Stream 3 in UL for AR are provided in our contribution R1-2107534 and the proposals captured below. |

# Summary of Contributions

Companies’ views on XR traffic model presented in contributions for RAN1#106-e are summarized as below [1]-[16].

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Huawei(2106456) | ***Proposal 1: For I/P-frame model for DL video, α = 2 is baseline. Other values can also be optionally evaluated.******Proposal 2: For I-stream and P-stream for DL video, RAN1 agrees on the following multiple combinations of (X, PDB) for evaluation. Other combinations can be optionally evaluated.****Table 2. (X, PDB) values for I-stream and P-stream for DL video*

|  |  |
| --- | --- |
| ***Index*** | ***(Packet success rate X%, PDB (ms)) of {I-stream, P-stream} in Option 1A and Option 1B*** |
| ***VR/AR*** | ***CG*** |
| *1* | *{(99.5, 10), (95, 10)}* | *{(99.5, 15), (95, 15)}* |
| *2* | *{(99, 15), (99, 9)}* | *{(99, 20), (99, 14)}* |
| *3* | *{(99, 10), (95, 10)}* | *{(99, 15), (95, 15)}* |
| *4* | *{(99, 15), (99, 10)}* | *{(99, 20), (99, 15)}* |
| *5* | *{(99, 15), (95, 10)}* | *{(99, 20), (95, 15)}* |

***The equivalent (PER, PDB) values for I-stream and P-stream, i.e. {(A, C), (B, D)} and {(E, G), (F, H)} are as follows****Table 3. (PER, PDB) values for I-stream and P-stream for DL video*

|  |  |
| --- | --- |
| ***Index*** | ***{(A, C), (B, D)} and {(E, G), (F, H)} of {I-stream, P-stream} in Option 1A and Option 1B*** |
| ***VR/AR*** | ***CG*** |
| *1* | *{(0.5, 10), (5, 10)}* | *{(0.5, 15), (5, 15)}* |
| *2* | *{(1, 15), (1, 9)}* | *{(1, 20), (1, 14)}* |
| *3* | *{(1, 10), (5, 10)}* | *{(1, 15), (5, 15)}* |
| *4* | *{(1, 15), (1, 10)}* | *{(1, 20), (1, 15)}* |
| *5* | *{(1, 15), (5, 10)}* | *{(1, 20), (5, 15)}* |

***Proposal 3:*** ***For evaluation of DL multiple streams, a UE is declared a satisfied UE if each stream meets the requirement that X (%) of packets are successfully delivered within a given air interface PDB.******Proposal 4: For I-stream and P-stream of AR UL video, i.e., Stream 2 and 3 in Option 4 for AR UL, RAN1 agrees on the following multiple combinations of (X, PDB) for evaluation. Other combinations can be optionally evaluated.****Table 4. (X, PDB) values for I-stream and P-stream for AR UL video*

|  |  |
| --- | --- |
| ***Index*** | ***(Packet success rate X%, PDB (ms)) of {I-stream, P-stream}*** |
| *1* | *{(99.5, 15), (95, 15)}* |
| *2* | *{(99, 15), (95, 15)}* |
| *3* | *{(99, 30), (99, 15)}* |
| *4* | *{(99.5, 30), (95, 30)}* |
| *5* | *{(99, 30), (95, 30)}* |

***The equivalent (PER, PDB) values for I-stream and P-stream, i.e. {(A, C), (B, D)} and {(E, G), (F, H)} are as follows****Table 5. (PER, PDB) values for I-stream and P-stream for AR UL video*

|  |  |
| --- | --- |
| ***Index*** | ***(PER, PDB (ms)) of {I-stream, P-stream}******{(A, C), (B, D)} or {(E, G), (F, H)}*** |
| *1* | *{(0.5, 15), (5, 15)}* |
| *2* | *{(1, 15), (5, 15)}* |
| *3* | *{(1, 30), (1, 15)}* |
| *4* | *{(0.5, 30), (5, 30)}* |
| *5* | *{(1, 30), (5, 30)}* |

 |
| ZTE(2106526) | [*Observation 1:* The average packet size ratio between I-slices and P-slices is ranging from 1.81 to 1.91 in VR2 configurations provided by SA4.](#_Toc16776)[*Observation 2:* The average packet size ratio between I-frames and P-frames is around 1.05 in VR2 configurations provided by SA4.](#_Toc25641)[*Observation 3:* The capacity performance is unchangeable whatever the average packet size ratio changes for the sliced-based traffic model.](#_Toc1947)[*Proposal 1:* The average packet size ratio is 2 as baseline and companies can optionally report the results with other ratios, including, e.g., 1.5 and 3.](#_Toc32272)[*Proposal 2:* Consider the parameters for I/P stream modeling in Table 5 as baseline.](#_Toc25156)***Table 5 Summary of parameters for I/P stream modeling***

|  |  |
| --- | --- |
| ***Application*** | ***AR/VR/CG*** |
| ***Two stream data*** | ***Stream #1: I-frame******Stream #2: P-frame*** |
| ***Option 1 Sliced-based*** | ***Option 2: Frame-based (GoP)*** |
| ***Structure*** | ***A frame consists of:******Number of Stream #1: 1******Number of Stream #2: 7*** | ***A GoP consists of:******Number of Stream #1: 1******Number of Stream #2: 7*** |
| ***Frame per second*** | ***Stream #1: 60FPS******Stream #2: 60FPS*** | ***Stream #1 + Stream #2 = 60FPS*** |
| ***Average packet size ratio*** | ***Stream #1 : Stream #2 = 2:1*** |
| ***(PSR, PDB)*** | ***AR/VR:******Stream #1: (99%, 20ms)******Stream #2: (90%, 20ms)******CG:Stream #1: (99%, 30ms)******Stream #2: (90%, 30ms)*** | ***AR/VR:******Stream #1: (99%, 10ms)******Stream #2: (90%, 10ms)******CG:******Stream #1: (99%, 15ms)******Stream #2: (90%, 15ms)*** |

[*Proposal 1:* The average packet size ratio is 2 as baseline and companies can optionally report the results with other ratios, including, e.g., 1.5 and 3.](#_Toc32272)[*Proposal 2:* Consider the parameters for I/P stream modeling in Table 5 as baseline.](#_Toc25156)***Table 5 Summary of parameters for I/P stream modeling***

|  |  |
| --- | --- |
| ***Application*** | ***AR/VR/CG*** |
| ***Two stream data*** | ***Stream #1: I-frame******Stream #2: P-frame*** |
| ***Option 1 Sliced-based*** | ***Option 2: Frame-based (GoP)*** |
| ***Structure*** | ***A frame consists of:******Number of Stream #1: 1******Number of Stream #2: 7*** | ***A GoP consists of:******Number of Stream #1: 1******Number of Stream #2: 7*** |
| ***Frame per second*** | ***Stream #1: 60FPS******Stream #2: 60FPS*** | ***Stream #1 + Stream #2 = 60FPS*** |
| ***Average packet size ratio*** | ***Stream #1 : Stream #2 = 2:1*** |
| ***(PSR, PDB)*** | ***AR/VR:******Stream #1: (99%, 20ms)******Stream #2: (90%, 20ms)******CG:Stream #1: (99%, 30ms)******Stream #2: (90%, 30ms)*** | ***AR/VR:******Stream #1: (99%, 10ms)******Stream #2: (90%, 10ms)******CG:******Stream #1: (99%, 15ms)******Stream #2: (90%, 15ms)*** |

[*Proposal 3:* Further discuss in RAN1 the parameters of FoV and non-FoV stream modeling for DL 360°video stream with parameters in Table 6 as a starting point.](#_Toc6910)***Table 6 Initial Parameters of FoV and non-FoV stream modeling***

|  |  |
| --- | --- |
| ***Application*** | ***VR1*** |
| ***Two Stream Data*** | ***Stream #1: FoV stream******Stream #2: Non-FoV stream*** |
| ***Option 1: sliced based traffic model*** | ***Option 2: Two separate streams*** |
| ***Structure*** | ***A frame consists of:******Stream #1: 1 (18 tiles)******Stream #2: 1*** | ***A Group of Tiles consist of:Stream #1: 18 tiles*** ***Stream #2: 1*** |
| ***Frame Per Second*** | ***Stream #1: 30FPS******Stream #2: 30FPS*** | ***Stream #1: 540 tiles per second******Stream #2: 30FPS*** |
| ***Data Rate*** | ***Stream #1: 12.78 Mbps******Stream #2: 8Mbps*** | ***Stream #1: 12.78Mbps (the aggregated data rate of the 18 tiles within a group of tiles)******Stream #2: 8Mbps*** |
| ***(PSR, PDB)*** | ***Stream #1: (99%, 20ms)******Stream #2: (90%, 20ms)*** | ***Stream #1: (99%, 10ms)******Stream #2: (90%, 10ms)*** |

 |
| vivo(2106629) | *Observation 1: For GOP-based traffic model with α=1.5, relaxing PDB of I-frame to 20ms can achieve similar capacity performance with single stream.**Observation 2: For GOP-based traffic model with α=3, it shows a significant performance degradation compared to single-stream traffic model.**Observation 3: Relaxing PDB of I-frame can improve capacity performance for GOP-based traffic model.**Observation 4: When I-frame and P-frame adopt the same PER and PDB values as single-stream traffic model, the capacity performance of slice-based traffic model is the same as that of single-stream traffic model, regardless of α=1.5 or α=3.**Observation 5: Relaxing PER of P-frame can improve capacity performance for slice-based traffic model.**Proposal 1: For GOP-based traffic model, =1.5 could be used as a starting point for evaluation purpose.**Proposal 2: For slice-based traffic model, =3 could be used as a starting point for evaluation purpose.****Proposal 3: Companies should report the PER and PDB values of I-frame and P-frame separately for two options of I/P-frame multiple stream modelling.****Proposal 4: 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 5: For GOP-based traffic model,* *the candidate values of E and F could be 1%, 5%, the candidate values of G and H could be 10ms, 15ms, 20ms.**Proposal 6: For slice-based traffic model, the candidate values of A and B could be 1%, 5%, the candidate values of C and D could be 5ms, 10ms.*  |
| Samsung(2106917) | *For AR/VR and CG: A=B=E=F=1%**For AR/VR: C=D=G=H=10ms and for CG: C=D=G=H=15ms* = 2 |
| CATT(2106949) | ***Proposal 1: The average size ratio α between one I-frame/slice and one P-frame/slice is 3.******Proposal 2: For the GOP-based encoding structure, the values of PER and PDB for the I-frame and P-frame are suggested as [PER\_I, PER\_P] = [1%, 10%] and [PDB\_I, PDB\_P] = [10ms, 10ms], respectively.******Proposal 3: For the slice-based encoding structure, the values of PER and PDB for the I-frame and P-frame are suggested as the same values, i.e. [PER\_I, PER\_P] = [10%, 10%] or [1%, 1%] and [PDB\_I, PDB\_P] = [10ms, 10ms], respectively*** |
| China Telecom(2107131) | ***Proposal 1:******Traffic arrival time offset among XR users per cell needs to adopt random offset with the random selection time in the [0 1/FPS] where FPS (Frame per second) is a frame refresh rate.******Proposal 2: Non-periodic modelling of traffic arrival time is recommended to be supported and FFS. For example, support 3GPP FTP model 3 as non-periodic model.******Proposal 3: Assume α = 1.5, 2 as baseline. Other values can be optionally evaluated.******Proposal 4: Assume [PER\_I, PER\_P] = [1 %, 10 %] or [0.5%, 5%] as baseline. Other values can be optionally evaluated, e.g., [0.1%, 5%].***Proposal 5: Assume [PDB\_I, PDB\_P] = [10ms, 10ms] or [15ms, 15ms] or [20ms, 20ms ] as baseline. Other values can be optionally evaluated. |
| OPPO(2107279) | Proposal 1: To model packet size correlation by using first-order autoregressive modeling, as described below.

|  |
| --- |
| Packet size modeling outputs:* {}: representing the generated packet sizes over the time, where each follows an identical truncated Gaussian distribution.

Packet size modeling parameter inputs: * {,}: respectively denote the minimum value and maximum value for the truncated Gaussian distribution, i.e., for all i.
* {µ,σ}: respectively denote the mean and standard deviation of the parent Gaussian distribution that is truncated. If {,} denote the mean and standard deviation for the (doubly) truncated Gaussian sample space {}, the relations between {µ,σ} and {,} are given by [3] (p35)

 , where , , , * ρ: denote the correlation coefficient between any pair of adjacent packet sizes in sample space {}, i.e., .

Packet size modeling procedure:* The procedure maintains an independent Gaussian random number generator X ~ N(µ=0,σ).
* Step-1: Repeatedly generate a random number x from generator X until . Assign , . Set n=1.
* Step-2: Generate a random number x from generator X. Assign and . If , increment n.
* Step-3: Loop back to Step-2.
 |

Proposal 2: RAN1 should agree upon the evaluation assumptions for two-stream traffic modeling in Option-2 (video + audio/data) |
| Qualcomm(2107374) | ***Observation 1**** ***The I and P frames/streams are generated periodically at edge server and displayed periodically at HDM/AR glasses, which means that the PDB values for I and P stream doesn’t need to be different.***

***Observation 2**** ***Due to the dependency of P frame on I frame, the effect of I frame loss is higher than that of P frame.***

***Observation 3**** ***Without actual measurement data, it is not clear how to map different PER requirements of I and P frames to final XR user experience.***

***Observation 4**** ***The XR capacity in single flow evaluation (or multi flow evaluation with PER\_I=PER\_P) cannot be directly compared with XR capacity in multi-flow evaluation (with PER\_I ≠ PER\_P) since the definition of satisfied UE is different, and it is not known how those conditions for UE satisfaction are mapped to final XR user experience.***

***Proposal 1******RAN1 do not directly compare XR capacity of single flow evaluation with that from multi-flow evaluation.*** |
| LGE(2107461) | ***Proposal 1: Companies to report the parameter values of*** α***, A, B, C, D, E, F, G, H, if they evaluate two streams of I-frame and P-frame for DL video stream (option 1).******Proposal 2: For optional two-stream DL traffic models, audio stream is aggregated with the data stream in Option 2*** |
| MediaTek(2107500) | ***Observation 1: The values of α, A, B, C, D, E, F, G, H in two streams traffic model of I-frame and P-frame for DL video stream are still FFS.******Proposal 1: Adopt α value to be 1.5 and 3.******Observation 2: Longer PDB for I-frames is needed as they have larger file size than P-frames (Di > Dp may provide capacity gain).******Proposal 2: Adopt [PDB\_I, PDB\_P] = [10ms, 10ms] and [17ms, 9ms].**** ***Equivalent to adopting (G, H) = (10, 10) and (17, 9)***

 ***It is noted that for GOP size K=8 in Option 1B, [PDB\_I, PDB\_P] = [10ms, 10ms] and [17ms, 9ms] provide the same average PDB.******Observation 3: Considering the larger importance of I-frame than P-frame, the PER of I-frame can be set equal or smaller than the P-frame.*** ***Proposal 3: Adopt [PER\_I, PER\_P] = [1%, 1%] and [0.5%, 5%]**** ***Equivalent to adopting (E, F) = (1, 1) and (0.5, 5)***

***Proposal 4: RAN1 to coordinate and cooperate with SA4 to construct a video quality evaluation block (as shown in the red block in Figure 2 below) based on statistical models used in RAN1 to evaluate the different QoS requirements and the performance enhancement for various RAN1 proposals.*** ***Proposal 5: 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.*** |
| InterDigital(2107534) | **Observation 1:** The UL traffic for AR are generally composed of video and non-video streams, where the non-video streams can consist of user action stream (e.g. . tracking data from sensors) and control data stream (e.g. protocol flow control, keep alive messages)**Observation 2:** The traffic models applied for AR/VR use cases (TR 38.824) can be used to model the user action traffic stream in UL for AR**Observation 3:** For AR, the UL traffic characteristics of aggregated video/media stream is significantly different than the pose/control stream**Observation 4:** For AR, due to the per-UE KPI requiring equal importance for all streams for meeting the respective X% and PDB and pairing of different streams with significant differences in traffic characteristcis, the non-video streams (e.g. pose/control, user actions) can have major impact on the capacity achievable Based on these observations, the following conclusions were made:**Proposal 1:** RAN1 uses user action traffic stream for UL evaluations of AR **Proposal 2:** RAN1 uses the following traffic model for the user action traffic stream (i.e. stream 3) in UL for AR* + - Packet arrival: Aperiodic with inter-packet arrival time (average) of 10ms
		- Packet size: 200 bytes
		- PDB: 4 ms

**Proposal 3:** RAN1 uses 3 traffic streams in UL for AR (i.e. aggregated video/media, pose/control and user actions) as baseline for capacity evaluations |
| Intel(2107616) | ***Observations-1 (from SA4 trace):**** ***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 VBR 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 (from SA4 trace):******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 (from SA4 trace):******We observe that*** *The average size ratio between one I-frame/slice and one P-frame/slice observed from SA4 traces is ~ 1.07– 1.90.* |
| Ericsson(2107629) | [Observation 1 Since traffic properties that greatly impact RAN performance have already been discarded, the value of providing more accurate modelling of DL video is very limited.](#_Toc79149097)Based on the discussion in the previous sections we propose the following:[Proposal 1 For A, B, C, D, E, F, G, and H, the same values are used as for DL video without I- and P-frame differentiation.](#_Toc79149098)This means that * For CG and AR/VR, A=B=E=F=1%
* For CG, C=D=G=H=15ms
* For AR/VR, C=D=G=H=10ms

[Proposal 2 The average size ratio between one I-frame/slice and one P-frame/slice is 3.](#_Toc79149099) |
| Apple(2107768) | For DL traffic model Option 2, the audio/data flow is modeled with:* A stream aggregating streams of audio and data
	+ Periodicity: 10ms
	+ Data rate: 0.756 Mbps/s or 1.12 Mbps
	+ Packet size: determined by periodicity and data rate
	+ PDB: 30ms
 |
| DOCOMO(2107886) | **Proposal 1:*** *Audio stream is aggregated with data stream in option 2 for modeling of DL multiple streams. The detailed traffic model for the aggregated streams can be as follows:*
	+ *Periodicity: 10ms*
	+ *Data rate: 0.756 Mbps/s or 1.12 Mbps*
	+ *Packet size: determined by periodicity and data rate*
	+ *PDB: 30 ms*
 |
| Xiaomi(2107905) | **Proposal 1: Send LS to SA4 to ask (X, PDB) requirement for packets associated to I-frames and P-frames** |

# List of contributions in RAN1 #106-e

1. R1-2106456 Traffic model for XR and Cloud Gaming Huawei, HiSilicon
2. [R1-2106526](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2106526.zip) Remaining Issues of XR Traffic Model ZTE, Sanechips
3. [R1-2106629](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2106629.zip) Remaining issues on traffic models of XR vivo
4. [R1-2106917](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2106917.zip) Traffic Models for XR Samsung
5. [R1-2106949](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2106949.zip) XR traffic model CATT
6. [R1-2107131](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107131.zip) Discussion on Traffic Model for XR/CG China Telecom
7. [R1-2107279](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107279.zip) Discussion on the XR traffic models for evaluation OPPO
8. [R1-2107374](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107374.zip) Remaining Issues on XR Traffic Models Qualcomm Incorporated
9. [R1-2107461](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107461.zip) Discussion on traffic models for XR evaluation LG Electronics
10. [R1-2107500](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107500.zip) Traffic Model for XR and CG MediaTek Inc.
11. [R1-2107534](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107534.zip) Discussion on UL traffic models for AR InterDigital, Inc.
12. [R1-2107616](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107616.zip) Traffic model for XR Intel Corporation
13. [R1-2107629](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107629.zip) Traffic model for XR Ericsson
14. [R1-2107768](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107768.zip) Remaining issues in XR traffic model Apple
15. [R1-2107886](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107886.zip) Discussion on traffic model for XR NTT DOCOMO, INC.
16. [R1-2107905](file:///C%3A%5CUsers%5Cyouns%5COneDrive%5CDocuments%5C3GPP%5CRAN1%20tdocs%5CTSGR1_106-e%5CDocs%5CR1-2107905.zip) Discussion on remaining issues of traffic Model for XR services Xiaomi