**3GPP TSG-RAN WG2 #113bis-e *R2-21xxxx***

**E-meeting, April 2021**

Agenda Item: 8.7.3

Source: Ericsson

Title: Summary of [602]

Document for: Discussion, Decision

# Introduction

This is to discuss the [602] as follows.

* [Post113bis-e][602][Relay] Definition of relay load criterion (Ericsson)

Scope: Collect definitions of the relay load criterion and downselect candidates. Whether to use relay load as a criterion will not be discussed in this scope.

Intended outcome: Report to next meeting

Deadline: Long

For rapporteur to have enough time drafting summary report, we would like to have the following two phases:

* Phase 1: collect companies’ views by 2021-05-06 1000 UTC
* Phase 2: rapporteur will finalize summary report based on inputs of phase 1 by 2021-05-10 1000 UTC

# Discussion

The necessity of relay load has been initially discussed in [1], company views were summarized as the following.

* 14/25 companies (ZTE, QC, Ericsson, Sony, IDC, Sharp, MDK, Xiaomi, Vivo, FH, Apple, Kyocera, Philips, Convida) believe load indication would be beneficial, or are open to considering load as an additional AS criteria in case something simple can be specified.
* A number of other companies indicated their concern for being able to specify something simple with the limited time available.
* From the options for defining load provided in the question, it seems those with the most support are a, b, and f. These also seem to be the ones which can lead to the simplest solution. Rapporteur suggests RAN2 continue to discuss (e.g. based on contribution) whether load is considered as an additional AS criteria while downscoping the options to the most popular (a, b, and f) only.

*Proposal 1-1: RAN2 continue to discuss further whether to consider load as an additional AS criteria, based on specific details of using number of PC5 connections/remote UEs served by the relay and/or resource pool usage/capacity at the relay UE.*

The above proposal was not agreed due to lacking sufficient support. Therefore, the following discussions base on the outcome from [1] to further collect definition of load criterion from companies and perform possible down-selection.

## Selection criterion

From Rapporteur’s understanding, the definition of relay load criterion shall fulfil the following conditions

1. Simple and easy to compute
2. Reflecting performance that a remote UE could achieve if served by the relay UE candidate
3. Small spec change
4. low signaling overhead

Bullet a) is important to make sure that introduction of relay load criterion will not increase relay UE’s implementation complexity.

Fulfilling bullet b, remote UE will be able to determine if a relay UE candidate can meet remote UE’s performance requirements of the relay traffic. Relay UE’s capabilities and remote UEs’ capabilities can be considered accordingly. With Bullet c) and d), RAN2 puts least design efforts for defining relay load criterion.

**Q1-1: do companies agree that the definition of relay load criterion shall fulfil the following requirements?**

1. Simple and easy to compute
2. Reflecting performance that a remote UE could achieve if served by the relay UE candidate
3. Small spec change
4. Low signaling overhead
5. Other (consistent interpretation of relay load with different capability of the Relay UE taken into account)
6. Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation

|  |  |  |
| --- | --- | --- |
| Company | Requirements of relay load criterion | Comments |
| Ericsson | a,b,c,d, |  |
| InterDigital | a, b, c, d |  |
| Qualcomm | A,b,c,d,e | For e), we think it is important because some relay UE can serve a large number of remote UEs while others can only serve a small number. The relay load metric should be able to reflect the capability difference of relay UE  Note that although rapporteur seemed to mention that b) reflects capability of relay UE, b) is not quite clear to us. That is why we make it clear in e) |
| vivo | a,b,c,d |  |
| OPPO | a,b,c,d,f |  |
| Nokia | a, b, c, d, e, f | Comment on b): Our understanding is that b) is about the "quality of service" that Remote UE can expect if the Relay UE is selected  Comment on e): Our understanding is that e) also covers that the load values provided by different UEs shall be comparable: a UE providing lower load value is expected to provide better service than a UE providing higher load value |
| Fraunhofer | a, b, c, d, e | For b), we agree with Nokia that this is about the required QoS (E2E) for the remote UE and our understanding is that existing mechanisms can be reused.  For e), we agree with QC in that it is an important aspect to take into account. However, we think it might help to initially define a common set of capabilities that the relay UE can support and make sure the relay load has a common interpretation across the set. |
| LG | a, b, c, d, e |  |
| Intel | a,b,c,d,f |  |
| Philips | a,b,c,d |  |
| Samsung | a,b,c,d,e |  |

## Definition of relay load criterion

Based on discussion outcome of clause 2.1, we list the most popular options which are indicated in [1] so that companies can do further evaluation and try to find a convergence among companies. In addition, companies are also welcomed to propose other options according to the requirements described in clause 2.1.

In this email discussion, Rapporteur refers the most popular options indicated in [1] (i.e., option a, b, and f) as Option 1, Option 2 and Option 3 in the below respectively.

**Option 1: Number of PC5 connections to Remote UEs currently being actively used for relaying**

**Option 2: Resource pool usage or capacity**

**Option 3: Number of remote UEs being served by the relay UE**

The above three options were not agreed in RAN2#113bis, mainly because that none of the options can reflect capabilities of relay UE and remote UEs which are being served by the relay UE. Therefore, the above options may not be able to indicate the real capacity or free bandwidth of relay UE candidate.

In addition, rapporteur would like to add an additional option, i.e.,

**Option 4** – **free bandwidth (or achievable bit rate) that relay UE can provide for relay traffic**.

In this option, a relay UE candidate can indicate how much bit rate or bandwidth a remote UE can achieve for its relay traffic if the remote UE connects to the relay UE candidate. This bit rate or bandwidth may be determined as the maximum bit rate/bandwidth of the relay UE candidate in Uu interface minus bit rate/bandwidth for relay traffic occupied by remote UEs which are being served by the relay UE in PC5 interfaces. This free bandwidth or achievable bit rate may be determined for UL relay traffic (i.e., from remote UE to gNB) and DL relay traffic (i.e., from gNB to remote UE) separately. In this option, relay UE candidate can estimate its maximum Uu bit rate/bandwidth based on implementation. It is feasible that relay UE candidate can perform estimation based on its radio channel quality or the historic UL grants or DL assignments. gNB may also provide assistance information (e.g. measure UL channel quality, and provide estimated UL bit rate to relay UE) accordingly.

In order to perform down-selection among the above options, Rapporteur would like to recommend companies to check whether each option can fulfill the requirements as described in clause 2.1.

**Q2-1: What requirements do companies believe that Option 1 is able to fulfil?**

1. Simple and easy to compute
2. Reflecting performance that a remote UE could achieve if served by the relay UE candidate
3. Small spec change
4. low signaling overhead
5. other (consistent interpretation of relay load with different capability of the Relay UE taken into account)
6. Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation

|  |  |  |
| --- | --- | --- |
| Company | Requirements of relay load criterion | Comments |
| Ericsson | a, c, d |  |
| InterDigital | a, c, d | Number of PC5-RRC connections should be easy to implement and signaling overhead should be limited. The relationship between number of PC5-RRC connections and the actual load on the relay may not be one-to-one. Some PC5-RRC connections may occupy a large number of resources (both sidelink resources and relay buffering capacity), while others would occupy a small amount of resources. |
| Qualcomm | A, c, d | It can’t reflect the capability difference of relay UE, although it is simple |
| vivo | a,c,d | The number of PC5-RRC connections is the most straightforward method for evaluation of load, but also a rough way as the capability of relay UE and the situation on each connection would be unknown. |
| OPPO | a,d,f | b: it should be the load status of each PC5 connection, which further multiplied by the number of PC5 connection of served remote UEs, that really reflecst the load status of the relay UE.  c: the definition of “being actively used for relaying” is not clear, so that RAN2 should firstly address the definition, in that case, quite much RAN2 spec impact will be caused. In fact, none of the options can achieve small spec impact since it includes not only cross-WG and cross-layer interaction (for discovery message design and AS/NAS layer interaction), and it requires further work on how for remote UE to behave w.r.t the value, i.e., how to define the AS layer criterion on the relay (re)selection. That is also means requirement-e cannot be achieved (i.e., “Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation”) |
| Nokia | a, c, d, f | This option does not meet b) as we think that the correlation between the number of PC5 connections and service quality that the Remote UE will experience is low; e.g.:  (i) if the Uu quality is low or the serving cell of the Relay UE is overloaded then Remote UE will not get a good a service even if the Relay UE has no other PC5 connection  (ii) if the other PC5 connections are not heavily used, and Uu quality is good then a Relay UE can easily provide good service for a Remote UEs  This option does not meet e) as it depends on the UE capability how many PC5 connections can be maintained |
| Fraunhofer | a, d | We also have the same concern as OPPO regarding the term ‘actively used for relaying’. For example, to be considered as actively relaying, should the relay UE calculate how often it is receiving data from the remote UE?  In addition, should we also only consider PC5-RRC connections?  For b), using Option 1 standalone might not be useful |
| LG | a, c, d | Option 1 is simple to compute and small spec change. But, it cannot reflect the amount of load in a relay UE. Even though relay UE has many numbers of active links with remote UE, there might just rare data transmission. Or Relay UE might have enough capacity to deal with many active links with remote UE. In this case, the relay UE cannot be said a loaded state.  We think Option 1 cannot represent the load of relay UE. |
| Intel | a,c, d | Option 1 is quite simple to advertise, however, we are not sure if it will accurately reflect the performance a remote UE could achieve. Depending on different scenarios and relay UE capability, it is not clear what reference the remote UE should utilize to evaluate and compare the candidate relay UEs. In general, we think that the Relay UE may perform discovery based on its own evaluation of supporting additional Remote UEs. |
| Philips | a,b,c,d,f | For b, in general one could say that the more Remote UE(s) actively making use of the same Relay UE, the less performance the Remote UE would be able to achieve |
| Samsung | a, d | This metric is simple to compute but we do not think that this metric represents the load of relay UE. With this number of PC5 connection, remote UE cannot decide whether it means high loaded relay or low loaded relay. |

**Q2-2: What requirements do companies believe that Option 2 is able to fulfil?**

1. Simple and easy to compute
2. Reflecting performance that a remote UE could achieve if served by the relay UE candidate
3. Small spec change
4. low signaling overhead
5. other (consistent interpretation of relay load with different capability of the Relay UE taken into account)
6. Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation

|  |  |  |
| --- | --- | --- |
| Company | Requirements of relay load criterion | Comments |
| Ericsson | a, c, d |  |
| InterDigital | a, b, c, d | If we stick to known metrics in SL, option 2 can be realized with either a CBR-like or CR-like metric (or average of this over time) which satisfies requirements a, c, and d. For requirement b, CR may be preferrable as it reflects the resource usage of the relay UE itself and has a direct relationship to how busy the relay is. CBR gives the overall load of the resource pool (which includes the usage of all other UEs including the relay). |
| Qualcomm | b) (others depend on exact form of metric) | It is not clear to us what is the form of metric for relay UE to calculate it. Proponent can provide a mathematical formula to help understanding. If without such detail, we can’t judge whether it is simple and what is spec impact |
| vivo | b | Option-2 can somehow reflect the performance in a perspective of resource usage but we agree with Qualcomm that it seems too general to consider what this metric can be like. |
| OPPO | See comment | Cannot find any relationship between the busy status of one specific resource pool and relay UE load. |
| Nokia | a, c, d, f | Our understanding is that the already available Channel Busy Ratio is meant here. This option does not meet b) and e), as CBR is a long time average (it does not have a strong correlation with the available service quality). Note also that it is only meaningful if the PC5 resource capacity is the bottleneck. |
| Fraunhofer | a, d | We also doubt the benefits of using this option to estimate relay UE load. In general, this might be a good indication for sidelink resource pool management. But how this could be tied to a specific relay UE(s) load is not clear to us |
| LG | See comment | We are not sure Option 2 can represent relay load. The past amount of resource usage cannot represent the load of relay UE. We have to consider the resource usage for the future. And also, we may have to consider priority and CBR. It seems not simple to compute the estimation of past and future resource usage. |
| Intel | See comment | We need further details of how this metric will be calculated. We may not be able to converge on this and it will be further effort to evaluate and converge on the relay load applicability. |
| Philips | a,b,c,d,f | Agree with InterDigital |
| Samsung | See comment | It is not clear how this metric represents relay load unless a resource pool is dedicatedly assigned to a relay UE. |

**Q2-3: What requirements do companies believe that Option 3 is able to fulfil?**

1. Simple and easy to compute
2. Reflecting performance that a remote UE could achieve if served by the relay UE candidate
3. Small spec change
4. low signaling overhead
5. other (consistent interpretation of relay load with different capability of the Relay UE taken into account)
6. Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation

|  |  |  |
| --- | --- | --- |
| Company | Requirements of relay load criterion | Comments |
| Ericsson | a, c, d |  |
| InterDigital | a, d | Similar to option 1, number of connected UEs may not directly indicate resource usage at the relay, since it depends on the services each remote UE uses. In addition, this may require larger spec impact to determine whether multiple PC5-RRC connections are associated with the same UE. |
| Qualcomm | A, c, d | It can’t reflect the capability difference of relay UE, although it is simple |
| vivo | a,d | The number of serving remote UE can be determined by the remote UE ID but this is a relatively rough method compared to option-1. It should anyway be the number of PC5-RRC connections and resource usage on each connection that matter. |
| OPPO | a,d | b: Similar as option 1, it should be the load status of one remote UE, associated with the number of remote UEs served by the relay UE, that really reflect the relay UE load.  c: Similar as option 1, it is not clear how to define “remote UEs served by the relay UE”, e.g. RRC idle remote UE but already selected/camped on one relay UE, whether it falls into the definition of “remote UEs served by the relay UE”. In fact, none of the options can achieve small spec impact since it includes not only cross-WG and cross-layer interaction (for discovery message design and AS/NAS layer interaction), and it requires further work on how for remote UE to behave w.r.t the value, i.e., how to define the AS layer criterion on the relay (re)selection.  f: the argument for c also means requirement-e cannot be achieved (i.e., “Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation”). Furthermore, according to the legacy PC5 connection mechanism, relay UE may not be aware of whether it has set up multiple PC5 links with one specific remote UE, in that case, relay UE cannot recognize the exact number of remote UEs being served in AS layer. Thus, SA2 needs to be consulted with. |
| Nokia | a, c, d, f | This option has similar issues as option 1, and thus our view is that it does not meet b) and e) |
| Fraunhofer | a, c, d | Similar to Option 1 using this option in a standalone manner might not be useful. In addition, potentially estimating the load just based on the number of UE(s) being served could be misleading. |
| LG | a, c, d | Same as the Option1 case. Option 3 is simple, but it cannot reflect the load of relay UE. It has a possibility of the wrong estimation for relay load. |
| Intel | a,c, d | Option 3, similar to option 1 is quite simple to advertise, however, we are not sure if it will accurately reflect the performance a remote UE could achieve using this input. Depending on different scenarios and relay UE capability, the number of UEs that could be supported by a relay UE may be quite different and it is not clear how relay UE can evaluate/compare among candidate relay UEs based on this |
| Philips | a,c,d,f | Less clear than Option 1, since Remote UE(s) may not be active |
| Samsung | a, d | Similar to Option 1, we do not think this metric represents the load in relay UE. |

**Q2-4: What requirements do companies believe that Option 4 is able to fulfil?**

1. Simple and easy to compute
2. Reflecting performance that a remote UE could achieve if served by the relay UE candidate
3. Small spec change
4. low signaling overhead
5. other (consistent interpretation of relay load with different capability of the Relay UE taken into account)
6. Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation

|  |  |  |
| --- | --- | --- |
| Company | Requirements of relay load criterion | Comments |
| Ericsson | a, b, c, d |  |
| InterDigital | b, d | While this achieves b, we see some challenges with how to compute the bit rate at the relay and how to specify it with minimal impacts. |
| Qualcomm | b) (at least it doesn’t fulfil a and e) | First, we think Option 4 is not quite clear what is the form of metric to reflect it? Proponent can provide a mathematical formula to help understanding. If without, we can’t judge whether it is simple and what is spec impact  Our first impression is that it is not easy for relay UE to calculate what is its free bandwidth (or achievable bit rate):   * It is up to relay UE implementation and hard to specify. * It may be power consuming for relay UE to calculate / update it time to time. * Absolute free bandwidth/bit rate may not be useful metric because a large absolute value may be a small fraction of relay UE’s capability, which is not preferred for relaying. |
| vivo | b | Similar view as Qualcomm. The formula should be first provided for companies to review, otherwise there is no need to consider a factor which is actually hard to compute. |
| OPPO | See comment | 1. Free bandwidth (achievable data rate) is not only related to UE capability and Uu load status, but also related to the available sidelink resource to be used, e.g. cr-limit. The definition is quite vague and is not straightforward. 2. So that there is no quite clear relationship between free bandwidth (achievable data rate) and relay UE load. 3. In fact, none of the options can achieve small spec impact since it includes not only cross-WG and cross-layer interaction (for discovery message design and AS/NAS layer interaction), and it requires further work on how for remote UE to behave w.r.t the value, i.e., how to define the AS layer criterion on the relay (re)selection. 4. the argument for c also means requirement-e cannot be achieved (i.e., “Able to be completed in RAN2 before the required deadline (RAN#92), i.e., in one meeting, and it needs to avoid inter-group consultation”) In addition, the computation method cannot be purely decided in RAN2, instead, RAN1 should be consulted with. |
| Nokia | d | The specification of the calculation of "free BW" is difficult, and the available free BW may not stable.  An option could be to leave the calculation for implementation, but in that case it will not meet criteria e). |
| Fraunhofer | b, d | More information is required to evaluate this criterion. Might be useful to combine it with Option 1 or Option 3 |
| LG | b | Our answer is similar to option 2. It’s not clear how to calculate the available bit rate. It has to be considered past and future available bit rate. And it can be different depends on CBR and priority. It seems not easy to compute. |
| Intel | b | Similar to option 2, we think that this will be extremely dynamic and might be difficult to quantize across different Relay UEs for a fair comparison (if as the rapporteur mentions, it is left to Relay UE implementation to come up with the bit rate, then, it will vary considerably). We still prefer to further understand this option and come back once its definition is fully available (however, it still is not clear whether we will be able to converge for this release). |
| Philips | b,d | If we consider this new option of achievable data rates, then we should consider latency as well, which may be equally (or even more) important for some UEs and some applications |
| Samsung | b | This metric may indicate available bandwidth at relay UE but we think that the calculation of free bandwidth/achievable bandwidth seems not simple to be standardized. |

Based on above questions, it is recommended to do down-selection of the options.

**Q2-5: According to the requirements as discussed in clause 2.1, which option do companies prefer for defining relay load criterion?**

**Option 1: Number of PC5 connections to Remote UEs currently being actively used for relaying**

**Option 2: Resource pool usage or capacity**

**Option 3: Number of remote UEs being served by the relay UE**

**Option 4: free bandwidth (or achievable bit rate) that relay UE can provide for relay traffic**

**option 5: Leave to UE implementation**

|  |  |  |
| --- | --- | --- |
| Company | Option | Comments |
| Ericsson | Option 4 |  |
| InterDigital | Option 2 | We prefer option 2, which in our opinion satisfies requirements a-d. |
| Qualcomm | Option 3 (if have to pick one among them) | We have concern on option 2/4 (at least for current unclear calculation metric) because both of them need relay UE to take its power to calculate capability/bandwidth timely and include it in discovery message (involving frequent NAS-AS interaction).  Option 1/3 is at least simple, although we are not sure whether they are useful. Similarly, including such time-variant info in discovery may incur frequent NAS-AS interaction. Between them, option 3 is more simple and less NAS-AS interaction. |
| vivo | Option 5 | From the analysis above we could tell that option-1 to option-4 all have their pros and cons and the main problem is none of them is a comprehensive way and can appropriately reflect the load overall.  Leaving to UE implementation would be another option which is flexible and can include all the options above. We prefer not to specify the concrete criterion which can be anyway hard to converge among companies, but instead to discuss about how the remote UE would know this relay load by relay UE, e.g., an indication of high/medium/low. |
| OPPO | None of above | Postpone to be discussed in Rel-18  we simply do not think this complicated issue which requires inter-WG / inter-layer interaction can be completed before the deadline of objective-1 in the WID, i.e., next RAN plenary. |
| Nokia | None of above | None of the options meet all the criteria. There are metrics that can be specified and calculated easily, but their relation to the actual service quality is vague.  It is not clear that the specification of a metric that have strong correlation with the service quality is feasible.  We think that this is not an essential item to provide relaying functionality. |
| Fraunhofer | Option 1/3 + Option 4 |  |
| LG | None of above | None of the options can meet the criteria. We think the methods of measuring load can be a UE implementation issue. A relay UE may be able to judge that the relay UE itself is loaded or not based on its implementation. If UE judges that the AS layer is over-loaded by itself (the judgment made by UE implementation), the UE can just inform it to its upper layer. Then the upper layer will decide how to act. |
| Intel | None (See comment) | As of now, we think that none of the options can fully and reliably capture the status at the Relay UE considering different applications/Relay Service Codes. In summary, options 1 and 3, although simple, may not support Remote UE’s decision making accurately. Option 4 may be quite dynamic and not fully indicative of the available bit rate when the Remote UE decides to perform relaying with the Relay UE. Moreover, as the Relay UE advertises Relay Service Code, it could be indicative of the services it supports and intrinsically, expected traffic.  Therefore, we agree with OPPO, Nokia and LG above that it may be difficult to converge on one parameter for this release due to limited time and potential complexity of introducing the discussed option(s). |
| Philips | Option 1, 2 or 4 | All 3 options are good starts for an initial proposal. We can further improve the proposal if we are granted more time to work on this issue in release 17 or 18, e.g. to consider also latency and/or other criteria such as the quality/channel state of the Uu link averaged over a period of time |
| Samsung | None of above | We do not think all the options are proper to represent the load in relay UE for remote UE. If relay load should be considered as a relay (re-)selection metric, then one bit indicator whether relay is loaded or not seems be enough for remote UE. How to decide the load at relay UE is up to relay UE implementation. |

1. xxxxx.

# Conclusion

We have the following proposal:

[Proposal 1 xxxxx.](#_Toc70023351)

3.1 For chair notes (proposal in priority order)

# Reference

[1] [R2-2104414](file:///C:\Users\mtk16923\Documents\3GPP%20Meetings\202104%20-%20RAN2_113bis-e,%20Online\Extracts\R2-2104414%20-%20%5b610%5d%5bRelay%5dAS_Criteria_reselection_summaryV2.docx) Summary of [AT113bis-e][610][Relay] AS criteria for relay (re)selection (InterDigital) InterDigital discussion

# Appendix