**3GPP T****SG-RAN WG3 Meeting #110-e R3-210997**

**Online, 25th January – 5th February 2020**

Agenda Item: 10.2.4

Source: Ericsson

Title: Summary of Discussion for SONMDT\_InterSystemLoad

Document for: Discussion, Decision

# Introduction

A Summary of Offline Discussions has been assigned to the topic of SONMDT\_InterSystemLoad.

The discussion has been summarised as follows in the meeting minutes:

**CB: # 1009\_SONMDT\_InterSystemLoad**

**CAC is enough? Other parameters (PRB utilization, Number of RRC connections, Number of active UEs, TNL Load, Hardware Load, etc.)?**

**Load information in HO signaling?**

**New threshold-based method?**

**May also discuss other issues based on papers submitted**

**Try to reach high-level agreements in the first phase, proceed to TPs in the second phase of the email discussion**

**(E/// - moderator)**

The following was agreed so far:

Introduce Inter System Load Balancing mechanisms on the basis of the solution available in E-UTRAN

Introduce Inter System Load Balancing by means of mechanisms that resemble or reuse the SON Configuration Transfer IE for the purpose of configuring load balancing metrics and reporting load balancing measurements

Use S1: eNB CONFIGURATION TRANSFER, S1: MME CONFIGURATION TRANSFER, NG: UL RAN CONFIGURATION TRANSFER and NG: DL RAN CONFIGURATION TRANSFER for the transfer of inter system load balancing via means of mechanisms that resemble or reuse the SON Configuration Transfer IEs. It is FFS whether further details on the signaling part need to be introduced

Adopt signaling of the Composite Available Capacity (Cell Capacity Class value and Capacity Value) for inter system MLB

Adoption of further MLB metrics is FFS

Event Based Reporting and Periodic Reporting (only in case specific conditions are met), are agreed to be supported for inter system MLB. The mechanism should avoid excessive signaling

Introduce a new mechanism for Inter System Status Request/Response/Update over NG: UL RAN CONFIGURATION TRANSFER and NG: DL RAN CONFIGURATION TRANSFER, via modification of the Inter-System SON Information IE

Introduce a new mechanism for Inter System Status Request/Response/Update over S1: UL RAN CONFIGURATION TRANSFER and S1: DL RAN CONFIGURATION TRANSFER, via reuse of the Inter-System SON Configuration Transfer IE

Support periodic inter system load reporting with periodicity not lower than 1000ms and threshold-based load reporting, subject to confirmation from CT

We do not support per slice load information for inter system load balancing in the current release

Support an explicitly signaled threshold configuration for inter system load information reporting; details are FFS

Agree to CAC encoding as defined in LTE, e.g. in TS36.413, as a starting point. Whether CAC is encoded according to the sender’s rules is FFS

Whether to support the Number of active UEs for inter system load balancing is FFS

*It is FFS whether to support signaling of PRB utilization for inter system load balancing in the current release;*

*It is FFS whether to support signaling of the Number of RRC connections for inter system load balancing in the current release;*

*It is FFS whether to support signaling of the TNL Available Capacity for inter system load balancing in the current release;*

*To be continued...*

# For the Chairman’s Notes

**Agreements after first round of discussions:**

* **Signalling of load information as part of HO messages is not supported in Rel17**

**Issues remaining pending after first round of discussions**

* **To continue discussions on CAC encoding, which also depend on development of other open issues (e.g. types of information reported)**
* **To discuss whether it is agreeable that CAC is the essential piece of information for inter system MLB and that enhancements to the inter system MLB solution may be achieved by means of exchanging PRB usage, number of RRC connections, number of active UEs**
  + **Companies proposing the inclusion of information other than CAC should describe how their signalling and node behaviour will work**
* **Two threshold mechanisms for inter system load balancing are proposed: range-based thresholds (legacy LTE) and explicit-thresholds (where each threshold can be flexibly selected). It is proposed to continue discussions on which threshold mechanism to follow.**

# Discussion

## Signalling mechanisms for Inter System Load Balancing

In past RAN3 meetings it was agreed that

Introduce a new mechanism for Inter System Status Request/Response/Update over NG: UL RAN CONFIGURATION TRANSFER and NG: DL RAN CONFIGURATION TRANSFER, via modification of the Inter-System SON Information IE

Introduce a new mechanism for Inter System Status Request/Response/Update over S1: UL RAN CONFIGURATION TRANSFER and S1: DL RAN CONFIGURATION TRANSFER, via reuse of the Inter-System SON Configuration Transfer IE

Support periodic inter system load reporting with periodicity not lower than 1000ms and threshold-based load reporting, subject to confirmation from CT

Support an explicitly signaled threshold configuration for inter system load information reporting; details are FFS

The above agreements point at a signalling solution for Inter System Load Balancing, where load information is either exchanged periodically, or where load information is exchanged at the occurrence of specific events.

In [1], a new proposal for signalling load information across systems is made. Namely, it is proposed to support the inclusion of load information as part of the inter system HO preparation signalling.

**Companies are invited to provide their view on whether Inter System load information exchange should happen also by means of including load information in the HO preparation signalling.**

|  |  |  |
| --- | --- | --- |
| **Company** | **Load Info within HO preparation: Yes/No** | **Comments** |
| Ericsson | No | The periodic and event-based reporting mechanisms enables all levels of LB information signalling granularity. Including LB info in HO preparation signalling may delay the HO preparation due to the need of deriving LB info before messages are signaled. Also, the procedure would be more error prone, in case any issue with Load Metrics is encountered at the receiver. |
| Nokia | Yes | The other forms of reporting require additional signaling. Including the information in the HO signaling could help. |
| Qualcomm | No | Periodic and event triggered load balancing gives us all the knobs to control the load reporting; don’t see much benefits in having another mechanism. |
| CMCC | No | Event-triggered and periodic reporting are enough. |
| China Unicom |  |  |
| Huawei | No | Similar view as Ericsson. Not needed and no benefit. In some sense the HO signaling is already providing the benefit to reject and thereby indicate an overloaded state. |
| ZTE | No | No obvious benefit in the HO signalling. Currently, the periodic and event triggered reporting are enough for us. |

**Conclusion: Signalling of load information as part of HO messages is not supported in Rel17**

## CAC encoding

One of the FFSs that remain to be solved is the following:

Agree to CAC encoding as defined in LTE, e.g. in TS36.413, as a starting point. Whether CAC is encoded according to the sender’s rules is FFS

In [2] and [3] it is proposed to adopt the rule of encoding CAC as per LTE specifications (i.e. as per TS36.413). In [5] it is proposed to report CAC according to the sender’s specifications of reference.

**Companies are invited to provide their views on the preferred way to encode the CAC for inter system load balancing**

|  |  |  |
| --- | --- | --- |
| **Company** | **Follow LTE Encoding/Encode as per senders specs** | **Comments** |
| Ericsson | Follow LTE Encoding | We would prefer to avoid impacts on LTE and therefore we would like to avoid that LTE needs to understand the new CAC structure supported in NR. |
| Nokia | Neutral | As long as CAC is used, it does not matter – it is defined in nearly identical way in LTE and in NR. |
| Qualcomm | Follow LTE Encoding | Same view as Ericsson. |
| CMCC | Sender’s rule | In NR, Capacity value is reported both per cell and per SSB.  As identified in our contribution, with the aid of measurement results obtained from UE and SSB CAC load reporting from gNB, eNB is able to making more precise and efficient decisions for inter-system MLB from E-UTRAN to NR.  In reply to E///’s comments: inter-system MLB is a R17 feature, we‘ll anyway upgrade our base stations if we would like to adopt new features. So the requirement that LTE needs to understand the NR CAC structure is not an obstacle. |
| China Unicom | Follow LTE encoding. | Same view as Ericsson. |
| Huawei | Follow LTE Encoding | Similar view as Ericsson. Better to keep LTE for simplicity. Also considering the event triggering |
| Samsung | Follow LTE Encoding |  |
| ZTE | No strong view | This issue may be related to 3.3. If we can agree to introduce some extra metrics besides CAC from NR to LTE. The encoding of CAC could follow the sender’s rule. |

**Conclusion: 5 companies support to Follow LTE Encoding; 1 company supports to follow the sender’s rules; 2 companies are neutral. It is proposed to continue discussions on CAC encoding, which also depend on development of other open issues**

## Inter System Load Metrics

A number of load metrics were discussed for inter system load balancing. It was already agreed to exchange the CAC, however the following FFS have been captured:

Whether to support the Number of active UEs for inter system load balancing is FFS

*It is FFS whether to support signaling of PRB utilization for inter system load balancing in the current release;*

*It is FFS whether to support signaling of the Number of RRC connections for inter system load balancing in the current release;*

*It is FFS whether to support signaling of the TNL Available Capacity for inter system load balancing in the current release;*

[1], [2] and [3] state that reporting the CAC is sufficient for inter system load balancing. In particular, the following observations were made concerning other load metrics:

* Even if PRB concept is valid for both LTE and NR, only a subcarrier spacing of 15 kHz is possible for LTE, which may be not the same as the one used in NR. This means that in general there is no common interpretation of PRB usage between LTE and NR
* The number of RRC connections can’t be interpreted if the total limit is not known. In light of the RRC Inactive feature, different implementation of RRC available states in LTE and NR provide different views in the two RATs
* The number of active UEs refers to UEs in RRC\_CONNECTED state for which there is data available for transmission or reception. This metric requires full knowledge of the resource structure of the sending node, as well as knowledge of the resource utilization. It is difficult to ensure such understanding between E-UTRAN and NG-RAN, hence we prefer to not use this metric for Inter-System MLB purposes.
* The TNL load information is different in LTE and in NR thus causing extra implementation burden; it should instead be included in the Composite Available Capacity

On the other hand, [5] proposes to signal between LTE and NR PRB usage, Number of RRC Connections and number of Active UEs. While [4] proposes that“The Number of active UEs, RRC Connections, TNL Capacity Indicator and PRB Utilization could be considered to transmit from gNB to eNB, while the Hardware Load, S1 TNL Load and PRB Utilization could be considered to transmit from eNB to gNB”

**In order to move forward, companies are invited to provide their view on the load metrics for inter system load balancing that are considered essential for Rel17**

|  |  |
| --- | --- |
| **Company** | **Comments** |
| Ericsson | CAC is the most important metric. We propose to use CAC for Rel17 |
| Nokia | CAC, same like Ericsson above. |
| Qualcomm | CAC should suffice. Signaling can also be simpler with limited amount of metrics. |
| CMCC | We support to introduce RRC connections, PRB utilizations and number of active UEs. In our opinion, CAC is important, but not enough.  In reply to arguments raised by [1]-[3]:  PRB usage: same as the comments we gave in 3.2, no common interpretation is not the obstacle. In fact, the scenarios we investigate for inter-system MLB is to choose from potential target eNBs by source gNB, and vice versa. Because all potential target eNBs uses the same interpretation on PRB usage in LTE, from gNB point of view, the load for these potential eNBs are comparable.  RRC connection: As specified in TS 28.552, RRC connection number only indicates the number of UE in RRC connected mode, where UE in RRC inactive mode is not counted. In addition, in latest 38413, a parameter called the number of stored inactive UE contexts which we believe is not exchanged on our specified interfaces yet (such parameter could be useful and we may introduce it later for MLB enhancement). So we don’t see any interpretation issue here. As the only metric that reflect control plane load, RRC connection is beneficial for inter-system load balancing, as identified by several companies last meeting. So it is highly recommended to be exchanged for inter-system MLB, and note that the extra overhead to introduce such a metric is not much.  Number of active UEs: similar reply as in PRB usage, please see comments above. Also note that the extra overhead to introduce such a metric is not much. |
| China Unicom | Besides CAC, agree with CMCC, PRB usage, RRC connection, number of active UEs are important metrics for load balancing. Other load information related with hardware could also be considered, such as TNL Capacity Indicator. From our view, the overhead for introducing such metric is not much and the benefits is obvious. |
| Huawei | CAC is most important. |
| Samsung | CAC is the most important metric. |
| ZTE | We agree that the CAC is the most important metric, but the other metrics should be not excluded, as the operators may need more information in the real deployment. |

**Conclusion: 6 companies think that the essential piece of information to be exchanged in inter system MLB is the CAC. 2 companies think that think that CAC and PRB usage, number of RRC connections, number of active UEs are essential.**

**It is proposed to agree that CAC is the essential piece of information for inter system MLB and that enhancements to the inter system MLB solution may be achieved by means of exchanging PRB usage, number of RRC connections, number of active UEs**

## Threshold structure for Inter System Load Balancing

In [3], [5] and [6] it is proposed to adopt a new threshold scheme different from the one use in LTE. This threshold scheme defines up to two thresholds, where the reporting event is triggered if the load metric goes below a threshold, or above a threshold or if the metric is in between thresholds.

In [2] a modification of the LTE threshold mechanism is proposed, where a subset of the range of values for the reported metric can be selected and where such subset is divided in equal parts by means of thresholds.

**Companies are invited to express their view on the threshold mechanisms preferred for inter system load balancing**

|  |  |
| --- | --- |
| **Company** | **Comments** |
| Ericsson | We prefer the mechanism with up to two thresholds defined in [3] because it allows to set thresholds in a flexible way (i.e. without keeping the rule of equal partitioning of the metrics values range). Such flexibility is needed to cope with both cases of load balancing (equal load across systems) and cases of load steering (higher loads in more capable systems). |
| Nokia | For simplicity, even threshold ranges are sufficient in Rel.17. More detailed information could be obtained if load info is appended to HO signalling. |
| Qualcomm | Threshold based load reporting gives more flexibility to give exact threshold values. We can define this method for event-triggered LB. |
| CMCC | Reuse even threshold method as the baseline, and to cope with the potential issue caused by introducing more load metrics than CAC, we prefer to use a CHOICE structure to indicate which load metric the event-triggered threshold setting is targeting for. |
| China Unicom | We are open to discuss two thresholds mechanism for inter system load balancing. |
| Huawei | We prefer the solution from LTE. It is a well known design and a reasonable starting point. If considered beneficial, we can also discuss adding the starting point as proposed by us. |
| Samsung | Prefer the solution from LTE. |
| ZTE | The explicit configured threshold structure(upper and lower thresholds) seems more straightforward. |

**Conclusion: 3 companies are in favour of explicit threshold based mechanisms. 3 companies are in favour of range based threshold mechanisms. 1 company proposes to adopt the range based threshold mechanism and the explicit threshold mechanism to cope with multiple load metrics reporting.**

**It is proposed to continue discussions on this topic.**

# Second Round of Discussions

## CAC and addition parameters to be exchanged

In section 3.3 it was discussed that the bear minimum needed to make inter system load balancing work is the signalling of the CAC. Namely, we can achieve a first implementation of the inter system load balancing by exchanging CAC between systems.

**Companies are invited to provide their view on whether a first implementation of the Inter System Load Balancing feature can be achieved by means of signalling the CAC between systems.**

|  |  |
| --- | --- |
| **Company** | **Comments** |
| Ericsson | We believe that a first implementation of inter system load balancing can be achieved by means of signaling of the CAC. We believe that enhancements to the load balancing mechanisms may be possible by means of exchange of other information, if this can be proven feasible. |
| Huawei | We think CAC is most important but we are also positive to continue discussing other measurements and analyze this as outlined in next section. As an example, we list some things to consider for a new measurement:  - do we set thresholds for both the new measurement and CAC, or only for one of them?  - do we signal the total nbr of available resources and a percentage of used resources for the new measurement? |
| CMCC | Not quite sure about the meaning of the first implementation.  However, it should be noted that we’ve agreed to adopt event-triggered reporting as well as periodic reporting for inter-system MLB.  1. Periodic reporting  During the online session, our feeling is that most of the concern on introducing more load metrics is focused on event-triggered reporting. And the additional load metrics can be introduced without too much difficulty for periodic reporting IF we follow the Sender’s Rule.  Of course the sender’s rule will require the receiver to understand the sender’s encoding; however, we don’t think it is the bottleneck of introducing additional metrics, since eNB/gNB will always be upgraded if new features are introduced in new release.  So our first proposal is to propose if we can **agree to introduce additional load metric (including PRB usage, RRC connections, Number of active UEs) for Periodic Reporting.**  2. Event-triggered reporting  To introduce additional load metrics for event-triggered reporting, we agree to include it as ‘to be continued’ with some outlined questions as indicated by the CB Moderator in Section 4.2 as below. In addition, more questions as indicated by HW’s comments above can also be captured. so that a simple and feasible solution on introducing additional load metrics for event-triggered reporting can be explained by more dedicated and comprehensive contribution for next meeting.  (Our primary thinking has been provided in our contribution in 0897, where we propose to use a CHOICE structure for threshold setting. The node requesting the event-triggered load reporting is able to choose which load metric it would like to focus on for inter-system MLB. It needs further modifications by considering the outlined questions raised during second round.) |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Details on extra parameters to be exchanged

During online discussions it was questioned how signalling of information other than the CAC may work. Namely:

* How can source and target understand the additional information, namely the impact on source and target should be outlined
* How would the signalling work with respect to reporting thresholds, e.g. ar ethe additional information reported when specific thresholds per information is reached? Are they reported all when only one threshold is met?

This discussion is worth taking with dedicated contributions and time, hence it is proposed to add this issue as “to be continued”

## Threshold mechanisms for inter system LB

We would like to point out that Load Balancing is not only about signalling load information when the cell load reaches high levels. Different policies may be in place for load balancing, for example:

* A policy by which cells of different systems are targeted to have the same load
* A policy by which cells of different systems are targeted to have different loads, e.g. due to a deliberately planned shift of load towards the better performing system.

Hence, one may want to set load thresholds not necessarily in uniform intervals like in LTE. As an example, one may want to set a threshold on an LTE cell when it reaches 50% load because load in excess of 50% should be shifted to/maintained on NR (due to operator’s policy). At the same time, one may want to set a threshold on NR at e.g. 90%, so to flag an overload and stop shifting of load.

With the LTE mechanism, the only way to set these thresholds (assuming to take the whole resource range into account) is to set thresholds every 10%, i.e. to set 9 thresholds.

These 9 thresholds will generate more than 4 times the amount of traffic than the originally needed two thresholds.

**Companies are invited to indicate which should be the essential mechanism for load threshold configuration in inter system load balancing, namely either the range based threshold mechanism or the explicit threshold mechanism**

|  |  |
| --- | --- |
| **Company** | **Comments** |
| Ericsson | We believe that the basic, simplest and most effective mechanism needed is the explicit threshold mechanism. We are open to add the range based mechanism if this is proved to be beneficial and feasible. |
| Huawei | The benefit of two level approach only exist when only two thresholds are needed and when the space between are not equal. The main difference is that with the range based scheme we are not restricted to two thresholds. Hence we prefer range based thresholds  But in line with the argumentation above, we think that adding a starting points to the range based scheme offers an efficient way to remove unnecessary thresholds from the low-load area.  For example, with the range method and by adding a starting point we can assign thresholds in the high range (e:g: 60%, 73% and 87% and skip the thresholds at the lower part. |
| CMCC | We are open to discuss both mechanisms. |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# Conclusion, Recommendations

# References

[1] R3-210078, Additional method for the inter-RAT load information exchange (Nokia, Nokia Shanghai Bell)

[2] R3-210400, (TP for SON BLCR for 38.413) Inter-System Load Balancing (Huawei)

[3] R3-210683, (TP for SON for TS 38.413, TS 38.300, TS 36.300): Inter-System Load Balancing BL CR (Ericsson)

[4] R3-210802, Further Discussion on Inter-system Load Balancing in NR (ZTE, China Telecom, China Unicom)

[5] R3-210897, More thoughts on inter-system load balancing (CMCC)

[6] R3-210804, (TP for SON BL CR 38.413) Inter-system Load Balancing (ZTE, China Telecom, China Unicom)