TSGR3#6-99988
(TSGW3#5(99)751)

TSG-RAN Working Group 3 meeting #6
Sophia Antipolis, France 24th – 27th August 1999

Agenda Item: 10.3 Procedure specifications (TS 25.413)
Source: **NTT DoCoMo, Nippon Telecommunications Consulting**
Title: Overload Control Procedure
Document for: Decision

References
[1] TS 25.413 RANAP specification version 1.1.1

1. Abstract

This contribution proposes to add the optional procedure to the current Overload control procedure in order to deal for the variety of situation.

2. Discussion

2.1 Congestion situation

Overload control procedure is used at abnormal situation. Especially when the traffic suddenly increases (e.g. concert ticket reservation, disaster), the quick and certain control is needed.

In the current specification, the traffic is reduced by one step on receipt of an OVERLOAD message. In this step by step procedure, though the congested situation is controlled with the simple method, if high reduction of traffic is required, it is necessary to send many messages until proper level is obtained. Since MMS may be connected to some RNCs, Overload message is sent to multiple requested nodes in the case of CN node congestion and the number of the message is increased. The fewer the messages are sent from the congested node, the lighter the load is. Also there would be some delay due to sending some OVERLOAD messages. It is desirable to shorten the time lag to achieve the proper restriction because the delay may cause fatal situation.

2.2 Traffic Control Level

To avoid the delay due to step by step procedure and the load of sending many messages, OVERLOAD message would indicate the level for traffic control (e.g. percentage, step value). In this case, even if the high level restriction was needed, the congested node could get the intended level by only one message for one requested node.

Also when requested node does not support the requested control level, the approximate level, which can be handled by requested node, can be applied.

This level indicator is considered to be optional to avoid the impact to the system without this alternative procedure.

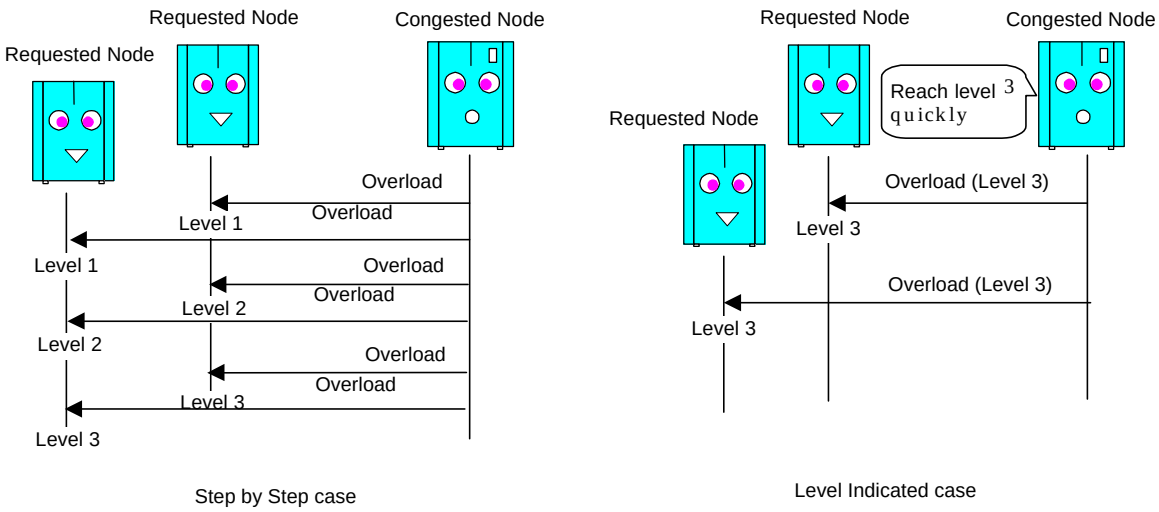


Figure 1.

2.3 Notification of traffic control completion

In the case of the introduction of traffic control level parameter, higher level restriction than necessary would be performed since congested node could not judge whether request message is reached or request level is too low.

For example, when a congested state does not break up after sending the Overload message due to the loss of the request, the congested node can not recognize the loss and may send the Overload message indicating higher level than previous request. If the second request is reached at the target node though the first request is proper level, higher level restriction is occurred. From operator's point of view, it is not desirable to restrict the call than necessary even in traffic congestion.

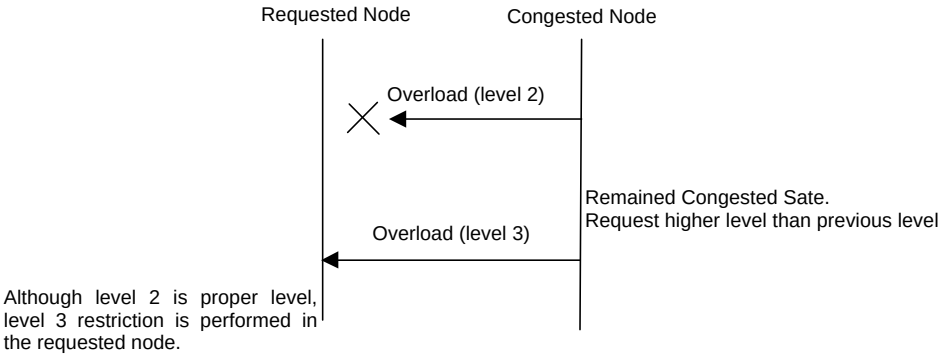


Figure 2

When the acknowledge for Overload message is introduced, the congested node could know the applied restriction level and decide whether more reduction is needed or not. Therefore it is expected to avoid the reduction of the traffic than

necessary and to perform efficient and precise control.

The acknowledge to Overload message should be sent back only when Traffic Control Level parameter is used.

3. Proposal

It has been proposed to add the Annex.1 to section 8.4 in TS 23.415 and Annex.2, 3 to proper sections.

Annex.1

8.6.1 Philosophy

The philosophy used is to stem the traffic at source with known effect on the service. The algorithm used is:

- On receipt of the first OVERLOAD message or signaling point congested information, the traffic is reduced by one step. At the same time, timers T(igOC)(T(igOR)) and T(inTC)(T(inTR)) are started. During T(igOC)(T(igOR)) all received overload messages or signaling point congested information are ignored in order not to reduce the traffic too rapidly. Reception of an OVERLOAD message or signaling point congested information after expiry of T(igOC)(T(igOR)) but still during T(inTC)(T(inTR)) , will decrease the traffic load by one more step, and restart T(igOC)(T(igOR)) and T(inTC)(T(inTR)).
- This step by step reduction of traffic is continued until maximum reduction is obtained by arriving at the last step. If T(inTC)(T(inTR)) expires (i.e. no OVERLOAD message or signaling point congested information is received during T(inTC)(T(inTR))) the traffic will be increased by one step and T(inTC)(T(inTR)) will be started, unless full load has been resumed.

  NOTE: Timers T(igOC) and T(inTC) are running in the CN whilst Timers T(igOR) and T(inTR) are running in the UTRAN.

- The number of steps and the method of reducing the load is considered to be an implementation specific function.
- It is possible for congested node to set the traffic control level in OVERLOAD message.
- Only when traffic control level is present in OVERLOAD message, traffic control should be performed with the level, which may be approximate level to the received traffic control level, and the Acknowledge message is sent as response.

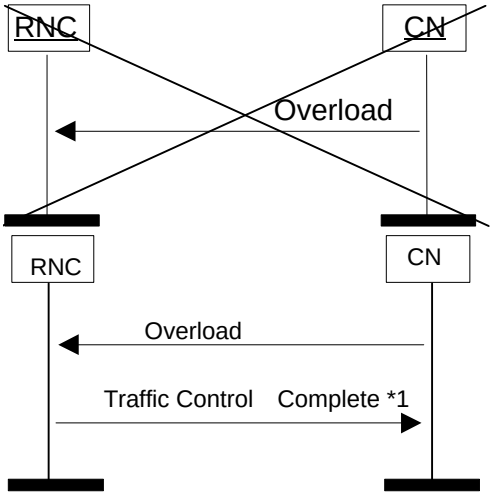
There may be other traffic control mechanisms from O and M activities occurring simultaneously.

8.6.2 Overload at the CN

The CN can indicate to the RNC that it is in a congested state by sending an OVERLOAD message. This is sent as a connectionless global message.

At the UTRAN receipt of this message causes the reduction of traffic to the CN node sending the message using the method described.

The signalling flow for Overoad at the CN is shown in Figure 26.



*1 : This message is used only when Overload message contains traffic level control parameter.

Figure 26. Overload at the CN.

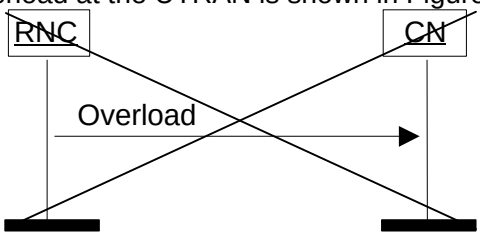
8.6.3 Overload at the UTRAN

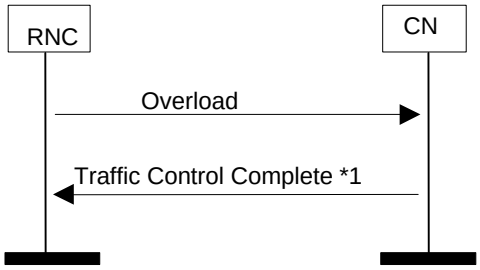
If the UTRAN is not capable to send signalling messages to the UE due to overloaded resources then the UTRAN sends an OVERLOAD message to the CN with the appropriate cause (Cause value: “overload in the capability to send signalling messages to the UE”).

If the UTRAN processing is overloaded then the RNC sends an OVERLOAD message with the Cause value: “processor overload”.

The CN originated traffic is reduced in accordance with the method described.

The signalling flow for Overload at the UTRAN is shown in Figure 277.





*1 : This message is used only when Overload message contains Traffic Control Level parameter.

Figure 27. Overload at the UTRAN.

Annex 2

OVERLOAD

INFORMATION ELEMENT	REFERENCE	TYPE
Message type		M
Traffic Control Level		O

Traffic Control Level can be ignored when RNC does not support this parameter..

TRAFFIC CONTROL COMPLETE

INFORMATION ELEMENT	REFERENCE	TYPE
Message Identifier		M

This message is used only when Traffic Control Level is present in OVERLOAD message.

Annex 3

TRAFFIC CONTROL LEVEL

8	7	6	5	4	3	2	1	
Parameter Identifier								1 (oct)
Length								2
Level								3