3GPP TSG RAN Meeting #27 **RP-050035**
Tokyo, Japan, 9 - 11 March 2005

3GPP TSG-RAN2 Meeting #45bis **Tdoc R2-050296**
Sophia Antipolis, France, 10-14 January 2005

Title: LS on ROHC testing
Response to: -
Release: Release 5 or 6
Work Item: RANimp-RABSE (RAB support enhancements)

Source: RAN2
To: T1
Cc: RAN

Contact Person:
Name: Håkan Olofsson
Tel. Number: +46 8 404 3590
E-mail Address: hakan.a.olofsson@ericsson.com

Attachments: R2-050140

---

1. Overall Description:

In response to concerns that the ROHC specification RFC3095 as such does not impose performance guarantees on implementations, RAN2 has decided to initiate work with the aim to define performance requirements and corresponding tests for ROHC UE implementations.

The approach is that requirements and test cases will be defined using ROHC expertise available to RAN2 in cooperation with testing expertise in T1. Specifications affected directly by this work will generally be under the responsibility of T1.

In the attached document R2-050140 (provided only for information), a first example of a possible testing approach is provided. RAN2 is still investigating this approach.

2. Actions:

To T1 group

ACTIONS:

1. *RAN2 kindly asks T1 to consider the general decision to perform a ROHC testing activity, and inform RAN2 if there are T1 testing aspects that RAN2 will have to take into account.*
2. *In particular, RAN2 kindly asks T1 to study the attached first outline of a testing approach, and give feedback on its feasibility from a T1 perspective.*
3. *Interested T1 experts are kindly invited to subscribe to the 3GPP ROHC email reflector (3GPP_TSG_RAN_WG2_ROHC), where it is foreseen that major parts of the work and discussions will be carried out.*

3. Date of Next TSG-RAN2 Meetings:

RAN2 #46	14th – 18th February 2005	Scottsdale, USA
RAN2 #46bis	4th – 8th April 2005	Beijing, China
RAN2 #47	9th -13th May 2005	EU, EU

3GPP TSG RAN WG2 meeting #45bis ***Tdoc R2-20050140***
Sophia Antipolis, France, 10-14 Jan 2005

Source: **Ericsson**

Title: **Performance testing of RObust Header Compression (ROHC)**

Agenda Item: *12.2*
Document for: *Discussion*

1 Introduction

A number of contributions to RAN2 have discussed and proposed the introduction of additional parameters to be signaled to the UE ROHC entity.

In this paper we propose an alternative approach that directly address what we believe is the intended goal behind this effort, namely to specify expected performance levels for implementations of RFC3095 through a suitable performance testing methodology.

2 Proposal

We propose to define a number of carefully selected performance requirements for UE compressor implementations of RFC3095. These requirements can provide a suitable level of performance and predictability for VoIP services, thus addressing the concerns expressed in RAN2.

The purpose of these performance requirements is to ensure that all conformant compressor implementations actually implement active ROHC compression, and for this to be verifiable using well-defined test cases. This will result in that the compression performance will meet the proper level of predictability, without preventing ROHC implementers from doing own optimizations and more advanced fine-tuning.

2.1 Overview of the Suggested Performance Testing Approach

For the purpose of performance testing of ROHC implementations, it is possible to define tests based on a set of different applications. However, at least initially, we suggest creating test cases for testing implementations performance using traffic patterns that corresponds to VoIP services only. It is important to understand that the test cases (along with associated parameters suggested herein) do not need to reflect a typical VoIP service in practice, but need only make sure that the factors that can impact compression efficiency are covered. This is because the purpose of the performance testing is to validate specific aspects of the ROHC implementations.

Our proposal is to use a test setup where one or more well-defined sequences of RTP packets are given as input to the compressor. Measurements are then performed on the output generated by the compressor. Some test cases may include injection of feedback to the compressor to emulate decompressor interactions.

We propose to measure the compressor performance level, and compare the results based on one or more well-defined thresholds for each of the test cases. Implementations that meet or exceed expectations with respect to these metrics would then be conformant to the performance requirements and known to properly implement active ROHC compression. This way, implementations that either implement robustness levels far above what is needed, implementations that take shortcuts resulting in less aggressive compression levels, as well as poor implementations due to other reasons will then fail to conform and can thus be identified.

Note that the performance requirements for the decompressor are discussed separately in section 4.

2.2 ROHC compressor performance testing

The compressor test setup would on the input side require one or more well-defined input packet streams (section 3.1). An additional input consisting of pre-defined decompressor feedback messages could complement the input stream. The purpose of this second input source is to provide the decompressor feedback necessary for each compression mode of operation, and to test the basic behaviour of the compressor in response to feedback (section 3.2). Each input packet stream defined, along with associated feedback input, would thus form a single test case.

The output is the header-compressed packet stream, over which the ROHC compression performance can be measured (section 3.3). To be declared 3GPP-comformant, an implementation of a ROHC compressor would then have to go below a specified threshold for the selected reference metric for the compressed header overhead.

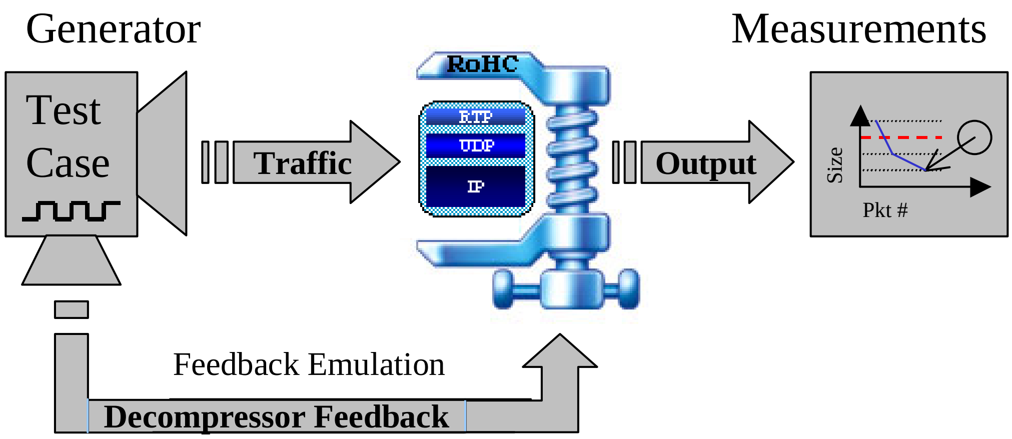

Figure 1: Conceptual Performance Test Setup for ROHC

The above figure represents the conceptual setup for the performance testing of a compressor. The feed-back emulation is defined from a ROHC perspective; it consists of ROHC decompressor feedback messages meant to test basic responses from the compressor for error recovery attempts. The generator performs the test cases, which consists of a stream of IP headers typical of VoIP. The output is the compressed headers generated from the input stimuli.

This contribution thus proposes that performance testing of the ROHC header compression mechanisms as described in RFC 3095 [2] be part of PDCP TTCN. This is pictured in Figure 2. A new test parameter should thus be added to control if ROHC performance testing is used as part of the PDCP testing. The PDCP Test Architecture is described in TS34.123 [3].

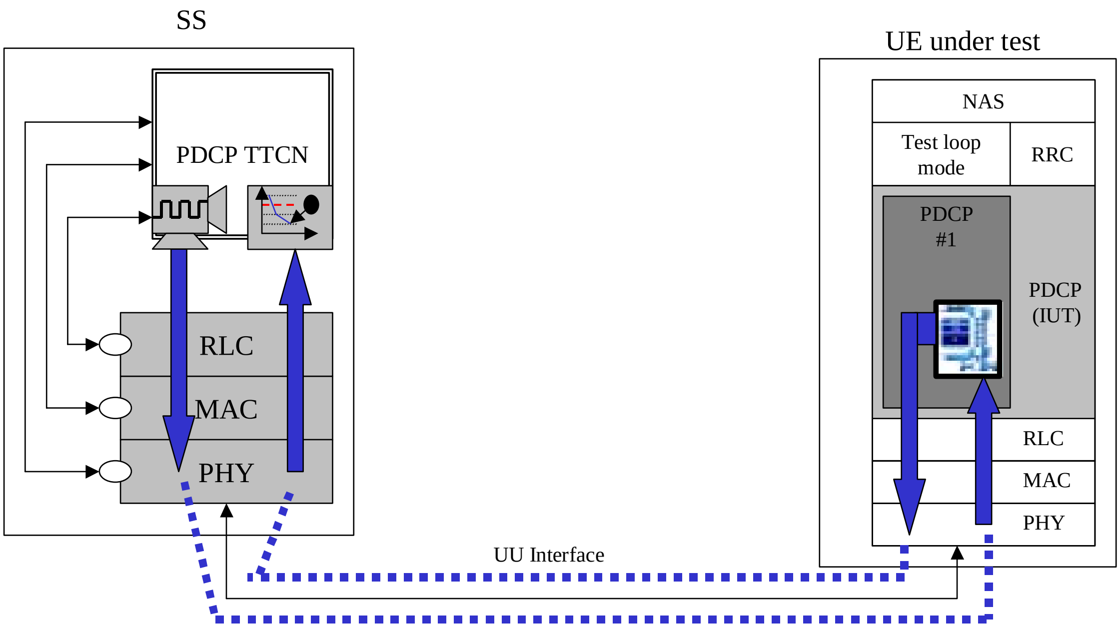

Figure 2: ROHC Performance Testing within the PDCP Test Architecture [3]

The ROHC performance testing should make use of uncompressed RTP/UDP/IP header packets of a pre-defined packet stream as well as pre-defined ROHC decompressor feedback, forming the test sequence input.

ROHC performance testing thus includes transmission of the aforementioned test sequence input and reception of compressed ROHC packets, as a part of the PDCP testing. It is meant only to test compression performance in terms of total overhead over the entire test sequence.

TS34.123 [3] defines a number of different input parameters as test parameter values necessary for PDCP testing. For ROHC performance testing, the test sequence input should be defined as part of those input parameters. Performance measurements should be performed inside the System Simulator (SS) [3]. The Test Case Generator should also reside in the SS. This means that all components creating the input stimuli and analyzing the output of the UE would reside in the SS.

ROHC performance testing can be added within the current PDCP test method as follows: once the correct PID values are assigned [3], the SS can start testing the UE compressor. The SS first sends the test sequence input to the UE. Then, the UE processes the received input: the UE compresses the received RTP/UDP/IP packet or processes ROHC feedback received. The UE sends back the output of its compressor over the test loop. The SS then performs all the measurements on the returned ROHC compressed packets. The UE must use a valid PID assignment and a properly configured PDCP PDU for transmission.

3 Characteristics of Input Sequences and Measurements

3.1 Uncompressed Input

The stream of packets input to the compressor, the *uncompressed input*, should be specified to include typical header field behaviour known to have a significant impact on the selection of the compressed packet format. This directly translates into compression efficiency. The following should be considered:

- some streams should be defined for IPv4, and some for IPv6;
- typical number of packets in the different input streams, based on a specific mean call hold time. For VoIP with 1 packet every 20ms, and considering for example a typical mean call hold time of most networks being in the range of 90-120 seconds, this means 4500-6000 packets without DTX, or 2250-3000 packets with 50% DTX.
- variations in TTL (for IPv4) and Hop Limit (for IPv6) variations;
- re-establishment of functions TS to SN, for example due to DTX. This requires a proper model for DTX. The extent of the skew in the SN-TS function to re-establish must also be defined.
- different IP-ID behaviours: sequential, sequential jump and random.

3.2 Feedback Input

3.2.1 Testing Compressor Basic Behaviour in Response to Feedback

We propose that this input be a specific sequence of feedback messages that is only meant to test basic compressor behavior as response to decompressor feedback. The feedback input does not need to be based on any specific link characteristics, or on emulation thereof. This is because the purpose of the proposed performance testing is to validate responses to feedback, and not to test robustness of the protocol itself.

More specifically, the basic behavior to be validated for the compressor includes the following, when receiving feedback:

- ACKs: the compressor should respond by sending second-order packets (when compression allows);
- NACKs: the compressor should respond by sending some first order packets to help the decompressor recover the dynamic part of the context (when compression allows);
- STATIC-NACKs: the compressor should send IR packets for complete recovery of the context.

Thus the artificial feedback messages would be constructed to trigger various compressor actions, and thereby verify proper handling of these feedback events. In addition, it may be useful to identify lazy compressors that would answer and handle NACKs as STATIC-NACKs, for example.

3.2.2 One Feedback Input Sequence per Modes of Operation

Because the decompressor uses different feedback patterns depending on the mode of operation, each test case should include one sequence of decompressor feedback corresponding to each of O-mode and R-mode.

3.2.3 Feedback Asynchronous to Compressor Output

There are two possible alternatives when defining the feedback input: the feedback messages can be either asynchronous or synchronized with the compressor output. Synchronization can be done using the sequence number field of ROHC feedback messages, which allows more advanced compressor implementation to be even more effective when answering error recovery requests from the decompressor.

However, we propose that the feedback input be defined as asynchronous messages, because it is the basic behaviour of the compressor that has the most impact. For example, the performance degradation is larger when IRs (e.g. instead of IR-DYN or UOR-2) would be systematically sent by a poor compressor implementation in response to a NACK from the decompressor than it would be if the compressor would systematically send IR-DYN, when some other slightly smaller first order packet could be sent would the compressor implement more advanced recovery algorithms.

In addition, performance testing of the compressor defined based on synchronized feedback input has implications in terms of the nature of the recovery algorithms that a compressor implements; much of this has been left open to implementations in RFC 3095. This would make assumptions related to how a compressor implementation performs additional recovery algorithms based on the sequence numbering information received in the feedback message when selecting the most suitable compressed header.

3.3 Performance Metrics

3.3.1 Metric used for the Proposed Testing

We propose to use the *target average compressed header size* as the performance metric for this testing as follow:

x% = compressed header size / uncompressed header size, for the entire test case

Alternatively, the total header overhead for the specific packet stream can also be used.

3.3.2 Measurements and Conformance

There are two possible interpretations of the compressor output with respect to the metric defined for the performance testing of the compressor: the measurement can be either:

1. a validation in terms of absolute performance in compression efficiency, i.e. the output is always smallest possible compressed header, or

2. a validation in terms of compression performance optimized for a specific bearer service, i.e. the output is always taking into account the finite number of packet sizes allowed by the bearer configuration.

In the latter alternative above, a compressor may be optimized to use the difference between the sizes allowed by the bearer services and the resulting compression to choose slightly larger compressed headers in order to improve robustness. Taking the set of allowed packet sizes into account for this testing can be realized by matching the size of the compressed output with the smallest allowed size available when calculating the metrics when the sizes are not matched by the compressor.

Depending on the definition of conformance, the *target average compressed header size* could be different for each case. There is a possibility that a compressor that might be specially optimized for the specific bearer service being tested may end up outside the conformance interval if the metric is defined based on absolute compression efficiency.

This kind of additional ROHC overhead to match the allowed packet sizes mandated by the bearer service in order to provide additional robustness can however be seen as an optimization per bearer service. We believe that such optimization should be left out of the performance testing, and that compressor implementation should be validated prior to any optimization to a specific bearer service.

We thus propose to validate the compressor performance based on the smallest possible compressed header size, i.e. alternative 1) above, since it is simpler to define, to test, it is well-understood from the RFC 3095 specifications and finally it is a more stringent requirement with respect to compression performance.

4 Decompressor Performance - Feedback

For the decompressor, the main concern of relevance is the interoperability of the various implementations, and this is normally handled through IOT (Inter Operability testing).

However, the manner in which a decompressor is implemented can also somewhat impact compression performance.

4.1 Overview of the Impact of Decompressor Feedback to Performance

More specifically, the potential impact relates to the feedback messages as generated by the decompressor implementation when error recovery is needed.

RFC3095 defines 3 types of feedback messages:

- an ACK is used to acknowledge successful decompression of a packet;
- a NACK indicates that the dynamic context of the decompressor is out-of-sync - at least one IR-DYN (or IR) packet must be sent by the compressor and successfully received by the decompressor in order to recover;
- a STATIC-NACK indicates that the static context is not valid - at least one IR packet must be sent by the compressor and received successfully by the decompressor to recover.

Negative feedback is generated for the purpose of error recovery. There is no single way for a decompressor implementation to realize feedback or local error recovery algorithms. Simpler implementations may have only few (or none at all) local repair algorithms and rely more than other implementations on the usage of feedback. Simpler implementations may also not implement the most efficient algorithms to identify the best actions required from the compressor in order to recover from repeated decompression failures and send e.g. a STATIC-NACK when a simple NACK would be sufficient. The simplest form of error recovery, for a passive decompressor, could at worse be to only make use of STATIC-NACKs as soon as decompression fails.

However, note that the potential impact on compression performance associated with the behaviour of the decompressor implementation can only be to a very limited extent. This is because RFC3095 was designed to be very robust to errors and packet losses over the link to minimize the risks of decompression failures: error recovery (and thus negative feedback) should thus occur infrequently for even the simplest decompressor implementation conforming to RFC3095.

In summary, the potential impacts of decompressor implementations to the overall compression efficiency are related to:

- the number and efficiency of the decompressor local repair mechanisms (i.e. repairs that do not require any intervention from the compressor);
- the algorithms triggering the different feedback messages.

4.2 Conclusions on Decompressor Performance Testing

ROHC performance requirements apply mainly to the compressor. For the decompressor, only interopeability is of relevance and this is handled by IOT.

The decompressor has very little means of impacting compression efficiency; i.e. NACKs comning from the decompressor have impact on the choice of packet the decompressor will make. In other words, the compressor may send larger packets as a result of the reception of a NACK from the decompressor. NACKs are sent only when decompression fails. However, since ROHC is by definition robust through the nature of its encoding methods (i.e. W-LSB encoding, etc.) and the use of a CRC carried within packets updating the context, the occurrence of NACKs should be very rare.

Once the performance testing has ensured that a 3GPP-conformant compressor reacts properly to decompressor feedback (including NACKs) as we are proposing, then performance testing of the decompressor is not really needed. Of course one could always think of a very nasty decompressor implementation that would send NACKs randomly or even when compression succeeds, but these kinds of extreme cases should be virtually inexistent in practice.

In summary, we currently believe that only interoperability is important for the decompressor and that Inter-Operability Testing (IOT) handles this. Would it be found during the test specification work that performance testing is also relevant for the decompressor, we believe that this would require only minimal additions.

5 References

[1] R2-050141: "Way forward on ROHC for VoIP in RAN2 specifications, Ericsson

[2] IETF RFC 3095: "RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed"

[3] TS34.123 User Equipment (UE) conformance specification; Part 3: Abstract Test Suite (ATS) (Release 1999)