Network Working Group
Request for Comments: 3409
Category: Informational

K. Svanbro
Ericsson

December 2002

Lower Layer Guidelines for Robust RTP/UDP/IP Header Compression

Status of this Memo

Copyright Notice

Abstract

This document describes lower layer guidelines for robust header compression (ROHC) and the requirements ROHC puts on lower layers. The purpose of this document is to support the incorporation of robust header compression algorithms, as specified in the ROHC working group, into different systems such as those specified by Third Generation Partnership Project (3GPP), 3GPP Project 2 (3GPP2), European Technical Standards Institute (ETSI), etc. This document covers only lower layer guidelines for compression of RTP/UDP/IP and UDP/IP headers as specified in [RFC3095]. Both general guidelines and guidelines specific for cellular systems are discussed in this document.

Table of Contents

1. Introduction

Almost all header compression algorithms [RFC1144, RFC2507, RFC2508] rely on some functionality from the underlying link layer. Headers (compressed or not) are expected to be delivered without any residual bit errors. IP length fields are inferred from link layer length fields. Packet type identification may be separated from the header compression scheme and performed at the underlying link layer. [RFC2509], for example, elaborates on how to incorporate IP header compression [RFC2507] in PPP [RFC1661].

It is important to be aware of such assumptions on required functionality from underlying layers when incorporating a header compression scheme into a system. The functionality required by a specific header compression scheme from lower layers may also be needed if incorporation of a header compression scheme is to be prepared without knowing the exact details of the final scheme.

This document describes lower layer guidelines for robust RTP/UDP/IP header compression [RFC3095] as specified by the ROHC working group. [RFC3095] will from this point be referenced to as ROHC. These guidelines should simplify incorporation of the robust header compression algorithms into cellular systems like those standardized by 3GPP, 3GPP2, ETSI, etc, and also into specific link layer protocols such as PPP. The document should also enable preparation of this incorporation without requiring detailed knowledge about the final header compression scheme. Relevant standardization groups standardizing link layers should, aided by this document, include required functionality in "their" link layers to support robust header compression.

Hence, this document clarifies the requirements ROHC put on lower layers, while the requirements on ROHC may be found in [RFC3096].

2. General guidelines

2.1. Error detection

All current header compression schemes [RFC1144, RFC2507, RFC2508] rely on lower layers to detect errors in (compressed) headers. This is usually done with link layer checksums covering at least the

compressed header. However, any error detecting mechanism may fail to detect some bit errors, which are usually called residual bit errors.

As for non-compressed IP packets, lower layers must provide similar error detection, at least for ROHC headers. ROHC has been designed not to increase the residual bit error rate (for reasonable residual error rates) compared to the case when no header compression is used. Headers passed up to the header decompressor should, however, have a residual bit error probability close to zero.

A ROHC decompressor might make use of packets with erroneous headers, even if they must be discarded. It is therefore recommended that such invalid packets are passed up to the decompressor instead of being discarded by lower layers, but the packet must then be accompanied with an error indication.

2.2. Inferred header field information

Some fields of the RTP/UDP/IP headers may be classified as inferred, that is their values are to be inferred from other values or from an underlying link layer. A ROHC decompressor requires that at least the following information can be inferred from any underlying link layer:

Packet Length (IPv4) / Payload Length (IPv6)

The received packet (with compressed header) length.

Length (UDP)

This field is redundant with the Packet Length (IPv4) or the Payload Length (IPv6) field.

In summary, all these fields relate to the length of the packet the compressed header is included in. These fields may thus be inferred by the decompressor if one packet length value is signaled from the link layer to the decompressor on a per packet basis. This packet length value should be the length of the received packet including the (compressed) header.

2.3. Handling of header size variations

It is desirable for many cellular link layer technologies that bit rate variations and thus packet size variations are minimized. However, there will always be some variation in compressed header sizes since there is a trade-off between header size variations and compression efficiency, and also due to events in the header flow and

on the channel. Variations in header sizes cause variations in packet sizes depending on variations of payload size. The following will only treat header size variations caused by ROHC and not packet size variations due to variations of payload size.

The link layer must in some manner support varying header sizes from 40 bytes (full RTP/UDP/IPv4 header) or 60 bytes (full RTP/UDP/IPv6) down to 1 byte for the minimal compressed header. It is likely that the small compressed headers dominate the flow of headers, and that the largest headers are sent rarely, e.g., only a few times in the initialization phase of the header compression scheme.

Header size variations and thus packet size variations depend on numerous factors. Unpredictable changes in the RTP, UDP or IP headers may cause compressed headers to momentarily increase in size, and header sizes may depend on packet loss rate at lower layers. Header size distributions depend also on the mode ROHC operates in. However, for e.g., a voice application, carried by RTP/UDP/IPv4, with a constant speech frame size and silence suppression, the following basic header size changes may be considered as typical:

In the very beginning of the speech session, the ROHC scheme is initialized by sending full headers called IR/DYN. These are the largest headers, with sizes depending basically on the IP-version. For IPv4 the size is approximately 40 bytes, and for IPv6 approximately 60 bytes. The IR/DYN headers are used typically during one round trip time, possible interleaved with compressed headers. After that, usually only compressed headers are sent. Compressed headers may vary in size from 1 byte up to several bytes. The smallest compressed headers are used when there is no unpredictable changes in header fields, typically during a talk spurt. In the beginning of a talk spurt, compressed header sizes may increase by one or a few bytes momentarily. Apart from increases due to new talk spurts, compressed headers may increase in size momentarily due to unpredictable changes in header fields.

ROHC provides some means to limit the amount of produced header sizes. In some cases a larger header than needed may be used to limit the number of header sizes used. Padding octets may also be used to fill up to a desired size. Chapter 6.3 (Implementation parameters) in [RFC3095] provides optional implementation parameters that make it possible to mandate how a ROHC implementation should operate, for instance to mandate how many header sizes that may be used.

2.4. Negotiation of header compression parameters

ROHC has some parameters that need to be configured in an initial setup phase. Which header compression profiles are allowed may have to be determined and also what kind of context identification (CID) mechanism to use.

The lower layers supporting ROHC should thus include mechanisms for negotiation of header compression parameters such as CID usage and header compression profile support. In certain environments, it might also be desirable to have mechanisms for re-negotiation of these parameters.

The negotiation must also make sure that compressor and decompressor use exactly the same profile, i.e. that the set of profiles available after negotiation must not include two profile identifiers with the same 8-bit LSB value.

For unidirectional links, this configuration might have to be performed out-of-band or a priori, and similar methods could of course also be used for bi-directional links if direct negotiation is not possible.

2.5. Demultiplexing of flows onto logical channels

In some cellular technologies flows are demultiplexed onto radio bearers suitable to the particular flows, i.e., onto logically separated channels. For instance, real-time flows such as voice and video may be carried on logically separated bearers. It is recommended that this kind of demultiplexing is done in the lower layers supporting robust header compression. By doing so, the need for context identification in the header compression scheme is reduced. If there is a one to one mapping between flow and logical channel, there is no need at all for context identification at the header compression level.

2.6. Packet type identification

Header compression schemes like [RFC2507, RFC2508] have relied on the underlying link layer to identify different kinds of headers by means of packet type identifiers on link layers. This kind of mechanism is not necessarily needed for ROHC since a ROHC packet type identifier is included in all compressed ROHC headers. Only if ROHC packets are to be mixed with other packets, such as packets compressed by other header compression schemes, must the link layer provide a packet type identifier. In such cases, or if ROHC is used on top of link layers already providing packet type identification, one (1) packet type identifier must be reserved for identification of ROHC packets. Thus,

only one ROHC packet type is needed to mix ROHC and e.g., RFC 2507 flows, or to support ROHC on links where packet type identifiers are already present.

2.7. Packet duplication

Exact duplications of one and the same packet may waste transmission resources and is in contradiction to compression. Even so, packet duplication may occur for various reasons. Packet duplication may also occur in different places along the path for a packet.

ROHC can handle packet duplication before the compressor but such packet duplications should be avoided for optimal compression efficiency. For correct ROHC operation, lower layers are not allowed to duplicate packets on the ROHC compressor-decompressor path.

2.8. Packet reordering

Lower layers between compressor and decompressor are assumed not to reorder packets, i.e., the decompressor must receive packets in the same order as the compressor sends them. ROHC handles, however, reordering before the compression point. That is, there is no assumption that the compressor will only receive packets in sequence.

2.9. Feedback packets

ROHC may operate in three different modes; Unidirectional mode (U-mode), bidirectional optimistic mode (O-mode) and bidirectional reliable mode (R-mode). A brief description of the modes can be found in chapter 4.4 of [RFC3095].

In U-mode it is not necessary to send any feedback from the decompressor to the compressor. O-mode and R-mode requires however that feedback messages from the decompressor to the compressor be sent. Feedback messages consist of small ROHC internal packets without any application payload. It is possible in ROHC to piggy-back feedback packets onto regular packets with ROHC compressed headers and payload, if there is ROHC type of compression in both the forward and reverse direction. However, this piggy-backing may not be desired or possible in some cases.

To support ROHC O-mode or R-mode operation, lower layers must provide transport of feedback packets from decompressor to compressor. If piggybacking of feedback packets is not used, lower layers must be able to handle feedback as small stand-alone packets. For optimal compression efficiency, feedback packets from the decompressor should be delivered as soon as possible to the compressor.

3. Cellular system specific guidelines

An important group of link layer technologies where robust header compression will be needed are future cellular systems, which may have a very large number of users in some years. The need for header compression is large in these kinds of systems to achieve spectrum efficiency. Hence, it is important that future cellular systems can efficiently incorporate the robust header compression scheme.

3.1. Handover procedures

One cellular specific property that may affect header compression is mobility and thus, handover (i.e., change of serving base station or radio network controller).

The main characteristics of handovers relevant for robust header compression are: the length of the longest packet loss event due to handover (i.e., the number of consecutive packet losses), and relocation of header compression context when necessary.

Depending on the location of the header compressor/decompressor in the radio access network and the type of handover, handover may or may not cause disruptions or packet loss events in the (compressed) header flow relevant for the header compression scheme. For instance, if soft handover is used and if the header compressor/decompressor reside above the combining point for soft handover, there will be no extra packet losses visible to the decompressor due to handover. In hard handovers, where packet loss events due to handover is introduced, the length of the longest consecutive packet loss is most relevant and thus should be minimized.

To maintain efficient ROHC operation, it should be ensured that handover events do not cause significant long events of consecutive packet loss. The term "significant" in this context relates to the kind of loss tolerable for the carried real-time application.

If hard handovers are performed, which may cause significant long events of consecutive packet loss, the radio access network should notify the compressor when such a handover has started and completed. The compressor could then be implemented to take proper actions and prevent consequences from such long loss events.

Cellular systems supporting robust header compression may have internal mechanisms for transferring the header compression context between nodes where contexts may reside, at or before handover. If no such mechanism for transferring header compression context between nodes is available, the contexts may be resynchronized by the header

compression scheme itself by means of a context refresh. The header compressor will then perform a new header compression initialization, e.g., by sending full headers. This will, however, introduce an increase in the average header size dependent on how often a transfer of context is needed. To reinitialize the context in such cases, the lower layers must indicate to the header compressor when a handover has occurred, so that it knows when to refresh the context. Chapter 6.3 (Implementation parameters) in [RFC3095] provides optional implementation parameters that make it possible to trigger e.g., a complete context refresh.

3.2. Unequal error detection (UED)

Section 3.1 states that ROHC requires error detection from lower layers for at least the compressed header. However, some cellular technologies may differentiate the amount of error detection for different parts of a packet. For instance, it could be possible to have a stronger error detection for the header part of a packet, if the application payload part of the packet is less sensitive to errors, e.g., some cellular types of speech codes.

ROHC does not require UED from lower layers, ROHC requires only an error detection mechanism that detects errors in at least the header part of the packet. Thus there is no requirement on lower layers to provide separate error detection for the header and payload part of a packet. However, overall performance may be increased if UED is used.

For example, if equal error detection is used in the form of one link layer checksum covering the entire packet including both header and payload part, any bit error will cause the packet to be discarded at the ROHC decompressor. It is not possible to distinguish between errors in the header and the payload part of the packet with this error detection mechanism and the ROHC decompressor must assume that the header is damaged, even if the bit error hit the payload part of the packet. If the header is assumed to be damaged, it is not possible to ensure correct decompression and that packet will thus be discarded. If the application is such that it tolerates some errors in the payload, it could have been better to deliver that packet to the application and let the application judge whether the payload was usable or not. Hence, with an unequal error detection scheme where it is possible to separate detection of errors in the header and payload part of a packet, more packets may be delivered to applications in some cases for the same lower layer error rates. The final benefit depends of course on the cost of UED for the radio interface and related protocols.

3.3. Unequal error protection (UEP)

Some cellular technologies can provide different error probabilities for different parts of a packet, unequal error protection (UEP). For instance, the lower layers may provide a stronger error protection for the header part of a packet compared to the payload part of the packet.

ROHC does not require UEP. UEP may be beneficial in some cases to reduce the error rate in ROHC headers, but only if it is possible to distinguish between errors in header and payload parts of a packet, i.e., only if unequal error detection (UED) is used. The benefit of UEP depends of course on the cost of UEP for the radio interface and related protocols.

4. IANA Considerations

A protocol which follows these guidelines, e.g., [RFC3095], will require the IANA to assign various numbers. This document by itself, however, does not require IANA involvement.

5. Security Considerations

A protocol which follows these guidelines, e.g., [RFC3095], must be able to compress packets containing IPSEC headers according to [RFC3096]. There may be other security aspects to consider in such protocols. This document by itself, however, does not add security risks.

6. References

[RFC1144] Jacobson, V., "Compressing TCP/IP Headers for Low-Speed Serial Links", RFC 1144, February 1990.

[RFC1661] Simpson, W., Ed., "The Point-To-Point Protocol (PPP)", STD 51, RFC 1661, July 1994.

[RFC2507] Degermark, M., Nordgren, B. and S. Pink, "IP Header Compression", RFC 2507, February 1999.

[RFC2508] Casner, S. and V. Jacobson, "Compressing IP/UDP/RTP Headers for Low-Speed Serial Links", RFC 2508, February 1999.

[RFC2509] Engan, M., Casner, S. and C. Bormann, "IP Header Compression over PPP", RFC 2509, February 1999.

[RFC3095] Borman, C., Burmeister, C., Degermark, M., Fukushima, H., Hannu, H., Jonsson, L-E., Hakenberg, R., Koren, T., Le, K., Liu, Z., Martensson, A., Miyazaki, A., Svanbro, K., Wiebke, T., Yoshimura, T. and H. Zheng, "Robust Header Compression (ROHC)", RFC 3095, July 2001.

[RFC3096] Degermark, M., "Requirements for robust IP/UDP/RTP header compression", RFC 3096, July 2001.

7. Author's Address

Krister Svanbro
Box 920
Ericsson AB
SE-971 28 Lulea, Sweden

Phone: +46 920 20 20 77
Fax: +46 920 20 20 99
EMail: krister.svanbro@ericsson.com

8. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.