

Network Working Group  
Internet-Draft  
Updates: [RFC1701](#), [RFC2784](#), [RFC2890](#) (if  
approved)  
Intended status: Informational  
Expires: January 30, 2017

F. Templin, Ed.  
Boeing Research & Technology  
July 29, 2016

**GRE Tunnel Level Fragmentation  
draft-templin-intarea-grefrag-04.txt**

Abstract

GRE tunnels use IP fragmentation for delivery packets that exceed the path MTU. However, IP fragmentation has been shown to be susceptible to reassembly errors at high data rates, and IP fragments may be unconditionally dropped by some middleboxes. This document therefore introduces GRE tunnel level fragmentation, which overcomes these issues.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 30, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	GRE Fragmentation Header . . . . .	<a href="#">3</a>
<a href="#">3.</a>	GRE Tunnel Level Fragmentation and Reassembly . . . . .	<a href="#">4</a>
<a href="#">4.</a>	IANA Considerations . . . . .	<a href="#">4</a>
<a href="#">5.</a>	Security Considerations . . . . .	<a href="#">5</a>
<a href="#">6.</a>	Implementation Status . . . . .	<a href="#">5</a>
<a href="#">7.</a>	Acknowledgements . . . . .	<a href="#">5</a>
<a href="#">8.</a>	References . . . . .	<a href="#">5</a>
<a href="#">8.1.</a>	Normative References . . . . .	<a href="#">5</a>
<a href="#">8.2.</a>	Informative References . . . . .	<a href="#">6</a>
	Author's Address . . . . .	<a href="#">6</a>

## [1.](#) Introduction

GRE is specified in the following RFCs:

[\[RFC1701\]](#)[\[RFC2784\]](#)[\[RFC2890\]](#)[\[RFC7676\]](#). GRE fragmentation considerations are further discussed in [\[RFC7588\]](#). In its current manifestation, GRE allows for fragmentation of the payload packet only if it is an IPv4 packet with the Don't Fragment (DF) bit set to 0. GRE also allows for IP fragmentation of the delivery packet, but IP fragmentation has been shown to be susceptible to reassembly errors at high data rates [\[RFC4963\]](#) and IP fragments may be unconditionally dropped by some middleboxes [\[I-D.taylor-v6ops-fragdrop\]](#).

A third option (introduced here) is for the GRE tunnel to perform "tunnel level" fragmentation and reassembly on the payload packet at the GRE layer. In this way, the ingress can fragment the payload packet (while treating the payload packet's headers as ordinary data) and encapsulate each fragment in a separate delivery header. The GRE header requires a new fragment header field to support this.

This tunnel level fragmentation method was first suggested in [Section 3.1.7 of \[RFC2764\]](#), and also appears in more recent works [\[I-D.templin-aerolink\]](#) [\[I-D.herbert-gue-fragmentation\]](#). [\[I-D.ietf-intarea-tunnels\]](#) provides the architectural background for tunnel fragmentation and reassembly.

Templin

Expires January 30, 2017

[Page 2]

## 2. GRE Fragmentation Header

Figure 1 shows the GRE header as specified in [RFC1701] but with a new optional "Fragment Header" and a new control bit "F":

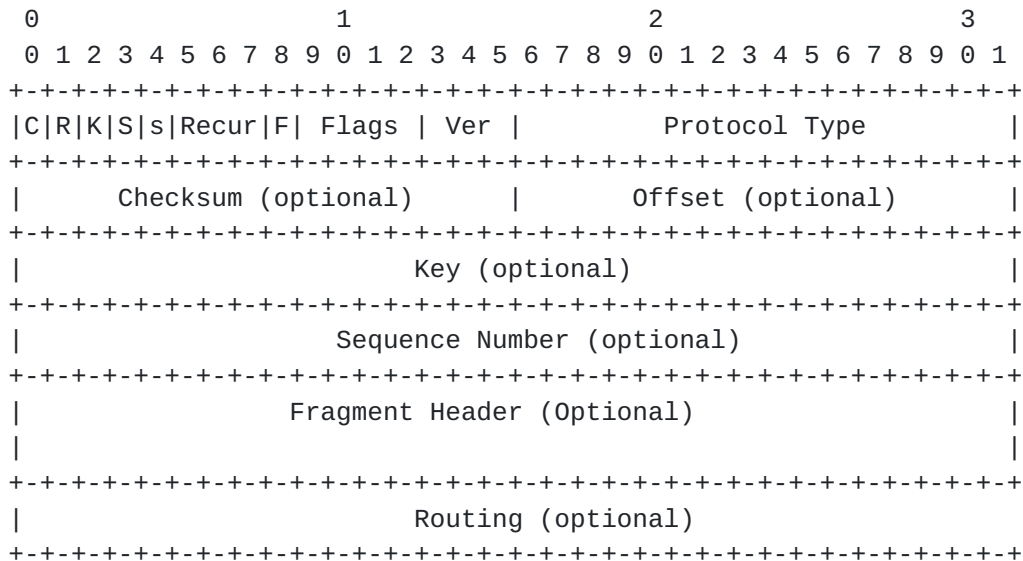


Figure 1: GRE Header with Fragment Header

In this format, when the "F" bit (i.e., bit 8) is set to 1 the GRE header includes a Fragment header formatted as follows:

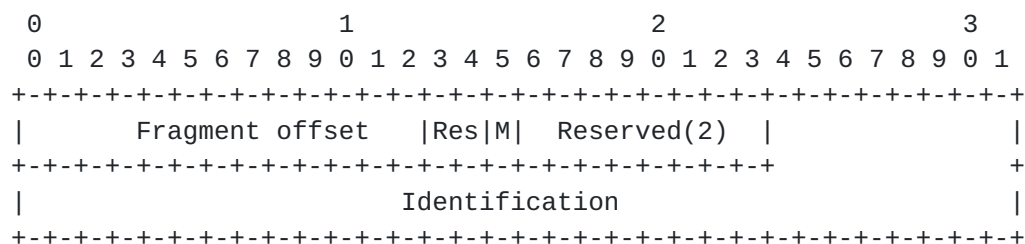


Figure 2: GRE Fragment Header Format

The fields of the option are:

- o Fragment offset: This field indicates where in the datagram this fragment belongs. The fragment offset is measured in units of 8 octets (64 bits). The first fragment has offset zero.
- o Res: Two bit reserved field. Must be set to zero for transmission. If set to non-zero in a received packet then the packet MUST be dropped.



- o M: More fragments bit. Set to 1 when there are more fragments following in the datagram, set to 0 for the last fragment.
- o Reserved(2): Eight bit reserved field. Must be set to zero for transmission. If set to non-zero in a received packet then the packet MUST be dropped.
- o Identification: 40 bits. Identifies fragments of a fragmented packet.

Note that these formats are the same as specified in [\[I-D.herbert-gue-fragmentation\]](#) with the exception that the Reserved(2) field replaces the "Original Type" field since the GRE header already includes a Protocol Type.

### **3. GRE Tunnel Level Fragmentation and Reassembly**

GRE tunnel level fragmentation treats the entire GRE payload packet (including the payload headers) as opaque data. The GRE tunnel ingress breaks the payload packet into N fragments and encapsulates each fragment in a separate GRE header and GRE delivery header. For the first fragment, the ingress writes the IEEE802 protocol number in the Protocol Type field the same as for any GRE packet. For other fragments, the ingress instead writes the length of the fragment in the Protocol Type field. This value MUST be no larger than 1500, which the egress will interpret as a length instead of a protocol type. (This implies that the maximum size for a non-initial fragment is 1500 bytes.) The GRE tunnel ingress then sends each fragment to the GRE tunnel egress.

When the GRE tunnel egress receives the fragments, it reassembles the GRE payload packet by concatenating the data portions of each fragment according to their offsets. In order to support this tunnel level fragmentation and reassembly procedure, the GRE tunnel ingress must know the maximum sized packet the GRE tunnel egress is capable of reassembling, i.e., the Maximum Reassembly Unit (MRU). In order to avoid interactions with Path MTU Discovery, the GRE tunnel egress MUST configure a minimum MRU of 1500 bytes plus the GRE delivery encapsulation overhead, and MAY configure a larger MRU.

### **4. IANA Considerations**

This document introduces no IANA considerations.



## **5. Security Considerations**

Security considerations for GRE apply also to this document.

## **6. Implementation Status**

The SEAL proposal uses tunnel level fragmentation in the same way as proposed here for GRE. Both SEAL and GRE fragmentation can be implemented through simple modifications of the widely-available, well understood and well-tested IP fragmentation code bases.

An implementation of SEAL fragmentation and reassembly has been published and is available at the following URL:

<http://linkupnetworks.org/seal/sealv2-1.0.tgz>

## **7. Acknowledgements**

The following are acknowledged for their helpful comments: Tom Herbert, Carlos Pignataro, Joe Touch.

## **8. References**

### **8.1. Normative References**

- [RFC0791] Postel, J., "Internet Protocol", STD 5, [RFC 791](#), DOI 10.17487/RFC0791, September 1981, <<http://www.rfc-editor.org/info/rfc791>>.
- [RFC1701] Hanks, S., Li, T., Farinacci, D., and P. Traina, "Generic Routing Encapsulation (GRE)", [RFC 1701](#), DOI 10.17487/RFC1701, October 1994, <<http://www.rfc-editor.org/info/rfc1701>>.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", [RFC 2460](#), DOI 10.17487/RFC2460, December 1998, <<http://www.rfc-editor.org/info/rfc2460>>.
- [RFC2764] Gleeson, B., Lin, A., Heinanen, J., Armitage, G., and A. Malis, "A Framework for IP Based Virtual Private Networks", [RFC 2764](#), DOI 10.17487/RFC2764, February 2000, <<http://www.rfc-editor.org/info/rfc2764>>.
- [RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", [RFC 2784](#), DOI 10.17487/RFC2784, March 2000, <<http://www.rfc-editor.org/info/rfc2784>>.





- [RFC2890] Dommetty, G., "Key and Sequence Number Extensions to GRE", [RFC 2890](https://www.rfc-editor.org/info/rfc2890), DOI 10.17487/RFC2890, September 2000, <<http://www.rfc-editor.org/info/rfc2890>>.
- [RFC7588] Bonica, R., Pignataro, C., and J. Touch, "A Widely Deployed Solution to the Generic Routing Encapsulation (GRE) Fragmentation Problem", [RFC 7588](https://www.rfc-editor.org/info/rfc7588), DOI 10.17487/RFC7588, July 2015, <<http://www.rfc-editor.org/info/rfc7588>>.
- [RFC7676] Pignataro, C., Bonica, R., and S. Krishnan, "IPv6 Support for Generic Routing Encapsulation (GRE)", [RFC 7676](https://www.rfc-editor.org/info/rfc7676), DOI 10.17487/RFC7676, October 2015, <<http://www.rfc-editor.org/info/rfc7676>>.

## 8.2. Informative References

- [I-D.herbert-gue-fragmentation]  
Herbert, T. and F. Templin, "Fragmentation option for Generic UDP Encapsulation", [draft-herbert-gue-fragmentation-02](#) (work in progress), October 2015.
- [I-D.ietf-intarea-tunnels]  
Touch, D. and W. Townsley, "IP Tunnels in the Internet Architecture", [draft-ietf-intarea-tunnels-03](#) (work in progress), July 2016.
- [I-D.templin-aerolink]  
Templin, F., "Asymmetric Extended Route Optimization (AERO)", [draft-templin-aerolink-70](#) (work in progress), July 2016.
- [RFC4963] Heffner, J., Mathis, M., and B. Chandler, "IPv4 Reassembly Errors at High Data Rates", [RFC 4963](https://www.rfc-editor.org/info/rfc4963), DOI 10.17487/RFC4963, July 2007, <<http://www.rfc-editor.org/info/rfc4963>>.

## Author's Address

Fred L. Templin (editor)  
Boeing Research & Technology  
P.O. Box 3707  
Seattle, WA 98124  
USA

Email: [fltemplin@acm.org](mailto:fltemplin@acm.org)

