Workgroup: loops
Internet-Draft:
draft-bormann-loops-geneve-binding-01
Published: 12 June 2020
Intended Status: Standards Track
Expires: 14 December 2020
Authors: C. Bormann
Universitaet Bremen TZI

Embedding LOOPS in Geneve

Abstract

LOOPS (Local Optimizations on Path Segments) aims to provide local in-network loss recovery. It can be used with tunneling protocols to efficiently recover lost packets on a single segment of an end-to-end path instead of leaving recovery to the end-to-end protocol, traversing the entire path.

[I-D.welzl-loops-gen-info] defines the information to be carried between LOOPS ingress and egress nodes in a generic way, giving a guideline on defining the common elements to embed LOOPS functions in various tunnel protocols. The present document specifies how to embed LOOPS in the overlay tunnel protocol chosen for the initial LOOPS specification, Geneve [I-D.ietf-nvo3-geneve].

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 https://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 14 December 2020.

Table of Contents

1. Introduction

LOOPS (Local Optimizations on Path Segments) aims to provide local in-network loss recovery. The LOOPS problems and opportunities draft [I-D.li-tsvwg-loops-problem-opportunities] illustrates some typical scenarios where LOOPS are applicable. One way to use LOOPS is to map it onto a tunnel protocol. The path segment on which LOOPS is applied then is a tunnel, which can be an existing one or created on purpose.

LOOPS allows the packet loss recovery to be performed over specific segments instead of end-to-end, enabling faster and more reliable data delivery. [I-D.welzl-loops-gen-info] defines the information to be carried between LOOPS ingress and egress nodes in a generic way, giving a guideline on defining the common elements to embed LOOPS functions in various tunnel protocols.

Geneve [I-D.ietf-nvo3-geneve] is an encapsulation protocol that can be used to create overlay tunnels. It defines an extensible TLV structure to carry so-called "tunnel options". The present document employs this flexibility, specifying how to embed LOOPS in Geneve. This specification covers the format and Geneve-specific procedures only: the actual LOOPS function and procedures are defined in [I-D.welzl-loops-gen-info].

LOOPS has two modes of loss recovery, retransmission and forward error correction (FEC). The current version of the present document covers retransmission only.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

This document makes use of the terminology defined in [I-D.welzl-loops-gen-info].

3. Geneve LOOPS Frame Format

Figure 1 shows the format of the Geneve Header and a single Geneve Option, as defined in [I-D.ietf-nvo3-geneve]. Geneve LOOPS defines a new Option class called LOOPS to carry LOOPS forward and backward information.

Geneve Header and Option:

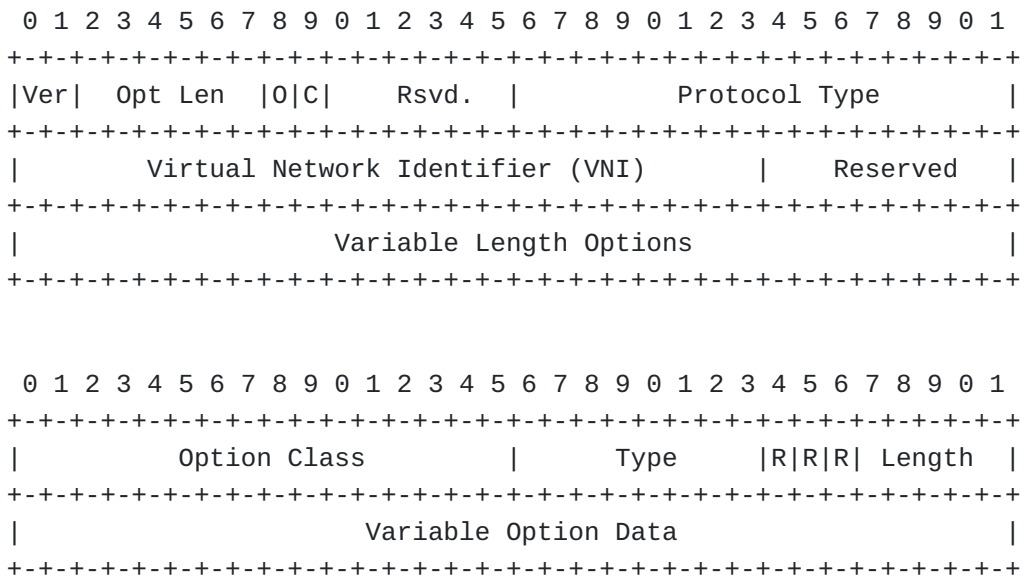

Figure 1: Geneve Header and Option Format

In the Geneve Option structure, a Geneve LOOPS option uses the following values:

*Option Class: TBD1 for LOOPS (see Section 5).

*Type: Based on the substructure already defined in Geneve, which uses bit 0 (the most significant bit) to indicate a critical option (see see Figure 2), LOOPS defines two type numbers: 0 for LOOPS retransmission mode, and 64 for FEC mode. The present document only addresses messages with LType=0.

TBD: Additional type numbers could be defined, possibly obviating the need for some of the flags in the current option structure.

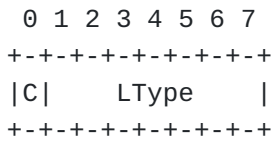

Figure 2: Type Field Format in Geneve LOOPS Option

*C: Critical bit as defined in [I-D.ietf-nvo3-geneve].

*LType: LOOPS Mode.

-0: Retransmission mode. In this mode, the LOOPS option format and operations follow this document.

-64: FEC mode

-Further mode values can be assigned in an IANA registry (see Section 5.2).

*Length: Length of Variable Option Data field, expressed in four byte multiples excluding the option header, ranging from 0 to 31. As the option header is another four bytes, the total length of the option in bytes is therefore 4 * (1 + Length), yielding a maximum total length of 128 bytes.

*Variable option data: consists of two parts, Flags and Flag Based Data, as shown in Figure 3.

-Flags: 16 bits, as described in next subsection. Some of the flags indicate the presence of additional data in the field of Flag Based Data.

-Flag Based Data: This field consists of one or multiple optional data blocks whose presence is indicated by the corresponding flag bits. Any remaining bytes needed to reach a multiple of four bytes are filled with zeroes.

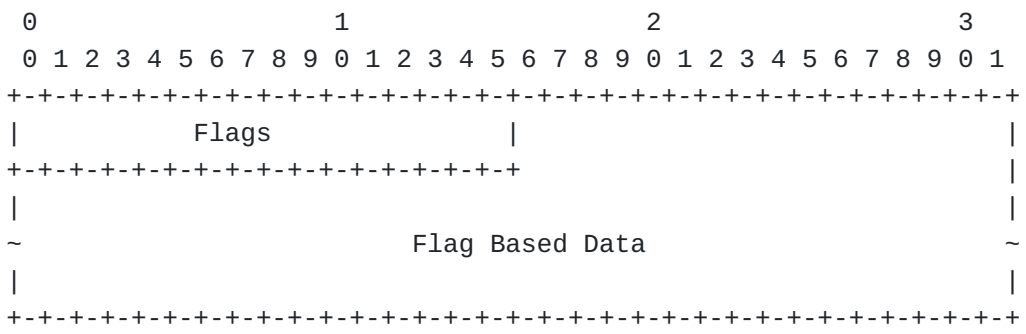

Figure 3: Variable Option Data Format in Geneve LOOPS Option

3.1. Flags and Flag Based Data

Flags for LOOPS Tunnel Options are defined in Figure 4. Some flags cause additional data blocks to occur in the Flags Based Data field. Those additional data blocks are placed in the order of the flags causing them.

Figure 4: Flags in Variable Option Data in Geneve LOOPS Option

A number of the flag bits are used on their own and do not cause carrying additional data:

- *I: Initial Packet Sequence Number (PSN) flag; may be set by the LOOPS ingress to notify the egress about using a new initial PSN.

- *R: Initial PSN Received flag; echo of I flag provided by the LOOPS egress.

- *D: ACK Desired flag; set by the LOOPS ingress if it wants the egress to generate an acknowledgement immediately upon receiving a particular packet.

These flag bits cause the addition of a single 32-bit number each:

- *S: PSN flag; indicates a PSN data block is carried in the Flag Based Data field. It must be set when a packet payload is present. It must not be set if the packet is a pure LOOPS ACK packet, i.e. when no payload is included in the packet.

*T: Timestamp flag. When set, it indicates a Timestamp data block is carried in the Flag Based Data field. TBD: Might want to have "timestamp" and "echo" fields of less or more than 4 bytes.

*E: Echoed Timestamp flag. When set, it indicates an Echoed Timestamp data block is carried in the Flag Based Data field.

*A: ACK number flag. When set, it indicates the presence of a Block 1 ACK information block.

*R: Reception time flag: May only be set if A is set. Indicates that an absolute reception time is given (Format TBD).

Finally, a single flag bit is defined that causes the addition of a variable-length block (therefore this flag is put as the least significant bit of Flags):

*B: Block 2 flag. When set, it indicates the presence a Block 2 ACK information block, with the following format: TBD copy over the structure we have in gen-info.

Acknowledgement information can be sent as a pure ACK packet without payload or piggybacked in a data packet.

4. Security Considerations

The security considerations of [I-D.welzl-loops-gen-info] and [I-D.ietf-nvo3-geneve] apply.

5. IANA Considerations

5.1. Geneve Option Class

IANA is requested to assign a new option class for LOOPS from the "Geneve Option Class" registry.

Option Class	Description
TBD1	LOOPS (Local Optimizations on Path Segments) [RFCthis]

Table 1

5.2. LOOPS Geneve Type Numbers

IANA is requested to create a registry for type numbers ("LType") as used in the TBD1 option class for LOOPS from the "Geneve Option Class" registry, with the following three columns:

Type Number: Integer between 0 and 127

Description: Short Description

Reference:
Reference to Specification

The initial contents of the registry is:

Type Number	Description	Reference
0	Retransmission mode	[RFCthis]
64	FEC mode	[RFCthis]

Table 2

(Registry policy TBD, probably Specification Required.)

6. References

6.1. Normative References

[I-D.welzl-loops-gen-info] Welzl, M. and C. Bormann, "LOOPS Generic Information Set", Work in Progress, Internet-Draft, draft-welzl-loops-gen-info-03, 9 March 2020, <http://www.ietf.org/internet-drafts/draft-welzl-loops-gen-info-03.txt>.

[I-D.ietf-nvo3-geneve] Gross, J., Ganga, I., and T. Sridhar, "Geneve: Generic Network Virtualization Encapsulation", Work in Progress, Internet-Draft, draft-ietf-nvo3-geneve-16, 7 March 2020, <http://www.ietf.org/internet-drafts/draft-ietf-nvo3-geneve-16.txt>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

6.2. Informative References

[I-D.li-tsvwg-loops-problem-opportunities]
Yizhou, L., Zhou, X., Boucadair, M., and J. Wang, "LOOPS (Localized Optimizations on Path Segments) Problem Statement and Opportunities for Network-Assisted Performance Enhancement", Work in Progress, Internet-Draft, draft-li-tsvwg-loops-problem-opportunities-04, 6 January 2020, <http://www.ietf.org/internet-drafts/draft-li-tsvwg-loops-problem-opportunities-04.txt>.

Acknowledgements

Sami Boutros provided some advice on the use of Geneve in this protocol binding.

Author's Address

Carsten Bormann
Universitaet Bremen TZI
Postfach 330440
Bremen D-28359
Germany

Phone: +49-421-218-63921
Email: cabo@tzi.org