Generic UDP Encapsulation (GUE) for Network Virtualization Overlay
draft-hy-nvo3-gue-4-nvo-03

Abstract

   This document describes network virtualization overlay encapsulation
   scheme by use of Generic UDP Encapsulation (GUE) [GUE]. It allocates
   one GUE optional flag and defines a 32 bit field for Virtual Network
   Identifier (VNID).

Status of This Document

   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 September 14, 2016.

Table of Contents

1. Introduction

Network Virtualization Overlay (NVO3) [RFC7365] provides a framework
for a virtual network solution over an IP network in a DC with
multi-tenant environment. Virtual network packets, i.e. tenant
packets, between any pair of Network Virtualization Edges (NVE) are
encapsulated at ingress NVE, sent from ingress NVE to egress NVE as
IP packets, and decapsulated at egress NVE. This is known as a
tunnel mechanism. This draft specifies use of Generic UDP
Encapsulation (GUE) [GUE] for NVO3 packet encapsulation.

GUE [GUE] as a generic UDP encapsulation provides several merits for
NVO3 encapsulation. Hence, underlay IP network treats it the same
was as other UDP applications, that are well supported by both IPv4
and IPv6 underlay networks. GUE provides strong security transport
options [GUE4SEC] that NVO3 can leverage. In addition, GUE supports
other options that NVO3 may use such as private data and
extensibility. In addition, GUE control flag can be used for NVO3
OAM message.

This document requests one flag (1 bit) from GUE optional flag field
for Network Virtualization Overlay (NVO3) indication and specifies a
32 bit field for virtual network identifier in GUE optional fields.
It describes use of GUE security options in NVO3.

2. Terminology

The terms defined in [GUE], [RFC7365] are used in this document.

   2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in [RFC2119].

3. Generic UDP Encapsulation (GUE) for NVO3

Generic UDP Encapsulation adds a 32 bit basic GUE header after UDP
header. GUE header contains some key fields that a UDP tunnel
application needs. These key fields are version, control message
indication (c), Header Length (HLen), and Protocol Type (or ctype).
It also contains some optional flags that are reserved for optional
features at a UDP tunnel.

   This document proposes to allocate one flag bit from GUE optional
   flags for the Network Virtualization Overlay (NVO3) and defines a 32
   bit field for NVO3 in GUE optional fields when the flag bit is set.
   GUE based NVO3 encapsulation format is shown in Figure 1.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |         Source port           |       Destination port        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |            Length             |           Checksum            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |Ver|C|  Hlen | Proto/ctype     |V|    Optional Flags           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |            Virtual Network Identifier (VNID)                  |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    ~                  Security, etc (optional)                     ~
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    ~                  Private Data (optional)                      ~
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```


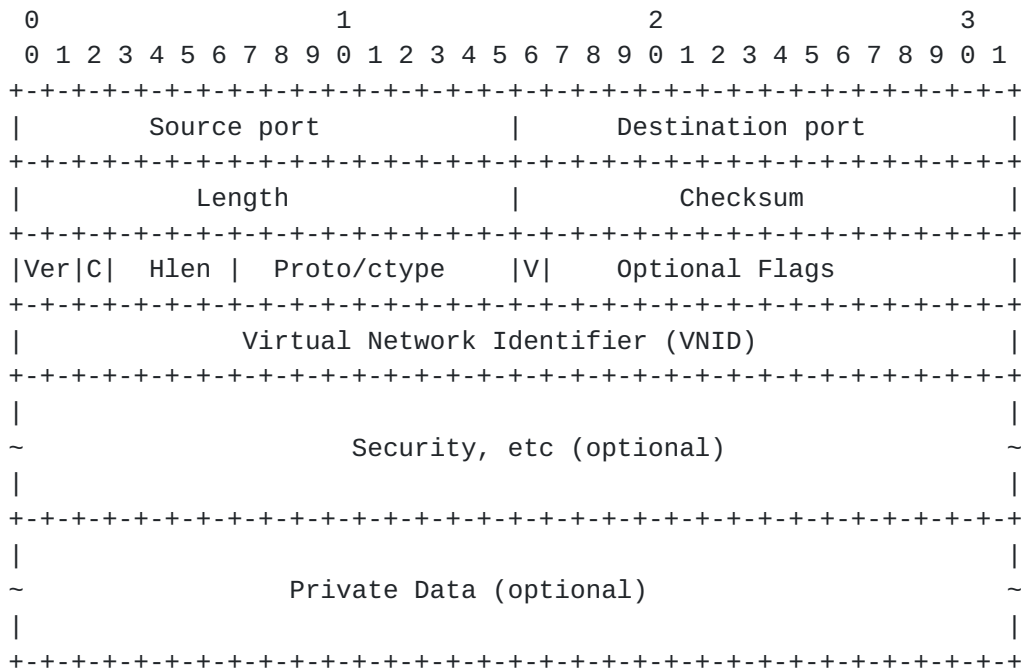   Figure 1 GUE based NVO3 Encapsulation Format

   o 'V': Virtualization flag. Indicates presence of the Virtual
      Network Identifier (VNID) field in GUE optional fields. This flag
      MUST be set when GUE is used for network virtualization overlay
      (NVO3).

   o Virtual Network ID (4 octets): a 32 bit field is used to identify
      a virtual network that the packet belongs to. This field MUST be
      present when 'V' virtualization flag is set; and MUST NOT present
      when 'V' flag is clear.

   NVO3 implementation may carry private data in the private data field.
   It must follow the rules specified in [GUE] when inserting private
   data in GUE header.

   NVO3 may allocate other flags and fields in GUE header for NVO3
   purpose and MUST follow the flag/field allocation rules specified in
   [GUE].

The usage of the key fields in the GUE header [GUE] for NVO3
encapsulation is described as below:

o Ver: Set to 0x0 to indicate version zero of GUE. Packets received
   by the decapsulator with non-zero version MUST be dropped.

o Control flag: When set, indicates that the packet contains a
   control message. OAM packets for the virtual network instance can
   be carried as a control message. NOV3 OAM packet format and
   mechanisms will be specified in a separated document.

o Hlen: length of (optional fields + private field (byte))/4.

o Proto/ctype: Contain the protocol of the encapsulated payload
   packet, i.e. next header or control message type (ctype) when
   Control flag is set. The next header begins at the offset
   provided by Hlen. For network virtualization, the payload
   protocol can be Ethernet, IPv4, IPv6, or 59 (NULL). The VNID can
   be used with ctype to direct control message for the VN layer.

UDP header field MUST be set per [GUE]. The checksum and length
implementation MUST be compliant with GUE implementation [GUE].

NVO3 can use GUE specified optional functions to improve the
transport such as GUE security option [GUESEC], GUE checksum option
[GUE4CSM], etc. When using a GUE specified option, NVO3
implementation MUST be compliant with the corresponding
specification.

4. Encapsulation/Decapsulation Operations

The network virtualization encapsulation by use of GUE applies to
both IPv4 and IPv6 underlay networks. The outer source and
destination IP addresses MUST be ingress NVE and egress NVE IP
addresses respectively. Ingress NVE adds UDP and GUE headers on the
payload packet with the required fields as described in Section 3.
NVE encapsulation and decapsulation process MUST be compliant with
GUE implementation specification [GUE]. If ingress and egress NVE
implement GUE options, they MUST be compliant with the corresponding
GUE option specification.

  4.1. Multi-Tenant Segregation

Ingress NVE MUST set option 'V' and insert Virtual Network
Identifier (VNID) into the corresponding option field when
encapsulating tenant packets. A GUE tunnel can carry the payload
packets that are from different tenant networks simultaneously.

Egress NVE MUST use the VNID in GUE header to identify the tenant
network that the payload packet is associated to and forward to the
packet to corresponding tenant network. All 32 bits can be used for
VNID.

4.2. Tenant Broadcast and Multicast Packets

If tenant packet is L2 broadcast/multicast, or L3 multicast packet,
depending on which multicast solution NV03 deploys [NVO3MFRWK], the
packet may be carried by a set of point-to-point GUE tunnels, or a
point-to-multipoint GUE tunnel. In the latter case, multicast IP
address is used as outer destination address.

The mapping of inner broadcast/multicast group to IP multicast group
can be manually configured or based on an algorithm, which is
outside the scope of this document.

4.3. Fragmentation

To gain the performance and simplification, NVO3 SHOULD avoid packet
fragmentation. Manual configuration or negotiation with tenant
systems can ensure that the MTU of the physical networks is greater
than or equal to the MTU of the encapsulated network plus GUE header.
It is strongly RECOMMENDED Path MTU Discovery [RFC1191] [RFC1981] to
be used by setting the DF bit in the IP header when GUE packets are
carried by IPv4 (this is the default with IPv6). In a case, it can't
avoid packet fragmentation; GUE fragmentation option can be used
[GUE4FRG].

4.4. GUE Header Security

NVO3 is expected to operate in multi-tenant environment, so security
is extreme important. Security can be provided by DC networking
and/or by NVO3. NVO3 can use GUE security options [GUE4SEC]. When
NVO3 use GUE security option, ingress NVE has to set the security
flag and insert a key value in the security field [GUE4SEC], egress
NVE has to validate the key prior to packet decapitation process. If
the key validation fails, the packet will be dropped [GUE4SEC]. The
key value used between ingress and egress NVE can be managed by NVA
or generated algorithm at NVEs. This mechanism will be described in
future version.

4.5. Tenant Packet Encryption

To prevent tenant packet from eavesdropping, tampering, or message
forgery, NVO3 can adopt GUE payload encryption mechanism. To encrypt
tenant packets, ingress NVE sets GUE payload transform flag and adds

32 bit payload transform field in GUE header. The payload type MUST
be filled at the payload transform field and the protocol field in
GUE base header MUST be set to 59 "No next header"[GUE4SEC]. Both
ingress NVE and egress NVE MUST implement the encryption mechanism
as described in [GUE4SEC].

5. IANA Considerations

The document request IANA to allocate the first bit in the registry
of GUE optional flag fields for Network Virtualization Flag and
register 32 bit field on GUE option field registry for Network
Virtualization Identifier (VNID).

6. Security Considerations

When Network Virtualization Edge (NVE) uses the UDP tunnel mechanism
specified in GUE [GUE], it faces the same security concern stated in
Section of Security Considerations in [GUE] and can leverage GUE
secure transport mechanisms [GUE4SEC] for secure transport over the
underlay IP network.

GUE provides two optional security functions. One is origin
authentication and integrity protection between encapsulator and
decapsulator, which protects Denial of Service (DoS) attacks;
another is GUE payload encryption, which prevents the payload from
eavesdropping, tampering, or message forgery. The two functions can
be used together or independently according to the deployment
environment. NVO3 virtual network identifier (VNID) is encoded in
GUE header that can be protected by origin authentication.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
          Requirement Levels", BCP 14, RFC2119, March 1997.

[RFC7365] Lasserre, M., el al, "Framework for Data Center (DC)
          Network Virtualization".

[GUE]   Herbert T. and Yong, L., "Generic UNP Encapsulation",
          draft-ietf-nvo3-gue, work in progress.

   [GUE4SEC] Yong, L., Herbert, T., "Generic UDP Encapsulation (GUE)
             for Secure Transport", draft-hy-gue-4-secure-transport,
             work in progress.

   [GUE4CSM] Herbert, T., "Checksum option for Generic UDP
             Encapsulation", draft-herbert-guecsum, work in progress.

   [GUE4FRG] Herbert, T., Templin, F., "Fragmentation option for
             Generic UDP Encapsulation", draft-herbert-gue-
             fragmentation, work in progress.

   7.2. Informative Reference

   [RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191,
             November 1990.

   [RFC1981] McCann, J., Deering, S., and J. Mogul, "Path MTU
              Discovery for IP version 6", RFC 1981, August 1996.

8. Authors' Addresses

   Lucy Yong
   Huawei USA
   5340 Legacy Dr.
   Plano, TX 75024
   US

   Email: lucy.yong@huawei.com

   Tom Herbert
   Google
   1600 Amphitheatre Parkway
   Mountain View, CA
   US

   Email: therbert@google.com

   Osama Zia
   Microsoft
   1 Microsoft Way
   Redmond, WA 98029
   US

   Email: osamaz@microsoft.com